

University of South Carolina Scholar Commons

Theses and Dissertations

2016

The Repeated Name Penalty and the Overt Pronoun Penalty in Japanese

Shinichi Shoji

University of South Carolina

Follow this and additional works at: <https://scholarcommons.sc.edu/etd>



Part of the [Linguistics Commons](#)

Recommended Citation

Shoji, S.(2016). *The Repeated Name Penalty and the Overt Pronoun Penalty in Japanese*. (Doctoral dissertation). Retrieved from <https://scholarcommons.sc.edu/etd/3745>

This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact dillarda@mailbox.sc.edu.

The Repeated Name Penalty and the Overt Pronoun Penalty in Japanese

by

Shinichi Shoji

Bachelor of Arts
Mie University, 1999

Master of Arts
Winthrop University, 2003

Master of Arts
University of South Carolina, 2014

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Linguistics

College of Arts and Sciences

University of South Carolina

2016

Accepted by:

Amit Almor, Major Professor

Stanley Dubinsky, Major Professor

Anne Bezuidenhout, Committee Member

Mineharu Nakayama, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

Acknowledgements

I thank the many people who have helped me complete my graduate work. Regarding my dissertation, I am especially grateful to my advisors, Amit Almor for his psycholinguistic input and statistical analyses, Stanley Dubinsky for his knowledge of Japanese syntax and semantics, and both of them for their instructions, encouragements and patience. I also thank my dissertation committee members, Anne Bezuidenhout and Mineharu Nakayama for their close reading and advice. In addition, I appreciate the help from Yuki Hirose, Koichi Sawasaki, Atsushi Fujimori, Katsuo Tamaoka and Etsuko Yoshida for me to collect data during my research trip in Japan. I also deeply thank all the professors from the linguistics program, including Mila Tasseva-Kurktchieva and Eric Holt for their constant and kind support and Paul Malovrh and Nina Moreno for being the committee members for my Master's degree. I would also like to thank my friends in the program, especially Wei Cheng, Peter Nelson and Burcu Gökgöz Kurt. I could not have survived this program without their support. My colleagues from the English Programs for Internationals and from the Japanese program have also been integral part of my life at this school. I thank my wife, Kazuko, for her support and my daughter, Karin, who was born when I was working on this dissertation. I appreciate the financial support from the USC Linguistics Program and the Mie International Exchange Foundation. Finally, I would like to thank all who have participated in my research.

Abstract

This research investigated the Repeated Name Penalty (RNP) and the Overt Pronoun Penalty (OPP) in Japanese. The RNP was first reported by Gordon, Grosz and Gilliom (1993), who observed that English sentences with repeated-name subject anaphors were read slower than sentences with overt-pronoun subjects when the antecedents were either the grammatical subject or the first-mentioned surface-initial noun phrase of the previous sentence. The OPP has been reported in studies of Spanish (Gelormini-Lezama & Almor, 2011) in which sentences with overt-pronoun subject anaphors were read slower than sentences with null-pronoun subject anaphors for subject antecedents.

A concern with the RNP and OPP is that, in most previously tested languages, the grammatical subject of a sentence is typically also its discourse topic. Thus, it remains unclear whether these effects are attributable to the anaphors' subject-hood or to their topic-hood. To address this question, the present study examined the RNP and the OPP in Japanese, a language that distinctly utilizes the topic-subject and non-topic-subject. In addition, while the existing studies controlled saliences of antecedents by grammatical status (i.e., subject vs. object), the present study utilized the information-structurally different antecedents, namely topic vs. non-topic (both were grammatical subjects in this dissertation study), and the non-linguistic aspects, empathy locus vs. non-empathy locus (both were grammatical subjects in this dissertation study).

The present study consisted of six self-paced sentence-by-sentence reading experiments with native Japanese speakers. The results showed that when antecedents were subjects or objects, the RNP and the OPP were detected only when anaphors were non-topic with the nominative-postposition, *ga*, but not topic anaphors with the topic-postposition, *wa*. The outcomes indicate that the topic-hood of anaphors activated immediate anaphoric interpretation. In contrast, when the antecedents were the topic or non-topic, the processing of anaphors with *wa* showed a marginal RNP, and anaphors with *ga* elicited no RNP. The results are attributed to the differences in salience between the antecedents and the function of the postpositions. Finally, although the different empathy statuses of antecedents elicited neither the RNP nor OPP, a closer examination of the results found a marginal OPP, likely reflecting the anti-logophoricity of overt pronouns in Japanese.

Overall, these results indicate that the topic-hood of anaphors should be considered in studies of the RNP and OPP and that the saliences of antecedents may be dependent on non-grammatical aspects such as discourse topic or empathy status in addition to grammatical status.

Table of Contents

Acknowledgements	ii
Abstract	iii
List of Tables	vi
List of Figures	ix
Chapter 1. Introduction	1
Chapter 2. Anaphors	7
Chapter 3. Adnominal Postpositions, <i>Wa</i> and <i>Ga</i>	41
Chapter 4. Antecedents	52
Chapter 5. Current Study	72
Chapter 6. Experiments 1 and 2	75
Chapter 7. Experiments 3 and 4	96
Chapter 8. Experiments 5 and 6	115
Chapter 9. General Discussion	143
Chapter 10. Conclusion	151
References	153
Appendix A. Readings and Meanings of Words Excluded from Corpus Search of <i>kare</i> ‘he’ and <i>kanojo</i> ‘she’	174

List of Tables

Table 2.1	The Repeated Name Penalty (Gordon et al., 1993)	8
Table 2.2	The Overt Pronoun Penalty (Gelormini-Lezama & Almor, 2011)	11
Table 2.3	The Givenness Hierarchy: Equivalent Anaphor Forms in English, Spanish and Japanese (Gundel et al., 1993, p. 284)	16
Table 2.4	Distinction in Pronouns (Shibatani, 1990, p. 371)	19
Table 2.5	Appearances of <i>Kare</i> ‘he’ and <i>Kanojo</i> ‘she’ in Books Published in 1971-2005	29
Table 2.6	Referents of <i>kare</i> in ‘Yamamoto is saying that <i>kare</i> will meet the division manager tomorrow.’ (<i>Kanno, 1997; Pimentel, 2014</i>).....	35
Table 5.1	Antecedents and Anaphors in Experiments 1-6	73
Table 6.1	Conditions and Example Items for Experiment 1 (RNP)	77
Table 6.2	Conditions and Example Items for Experiment 2 (OPP)	78
Table 6.3	Omnibus RNP Analyses for Experiment 1	83
Table 6.4	Separate RNP Analysis for Experiment 1: Null Pronoun vs. Repeated-name Non-topic- <i>ga</i>	84
Table 6.5	Separate RNP Analysis for Experiment 1: Null Pronoun vs. Repeated-name Topic- <i>wa</i>	85
Table 6.6	Simple Subject vs. Object Antecedent Effects for Anaphor Form in Experiment 1	86
Table 6.7	Omnibus OPP Analyses for Experiment 2	88
Table 6.8	Separate OPP Analysis for Experiment 2: Null Pronoun vs. Overt-pronoun Non-topic- <i>ga</i>	89

Table 6.9	Separate OPP Analysis for Experiment 2: Null Pronoun vs. Overt-pronoun Topic- <i>wa</i>	90
Table 6.10	Simple Subject vs. Object Antecedent Effects for Anaphor Form in Experiment 2	91
Table 7.1	Conditions and Example Items for Experiment 3 (RNP)	98
Table 7.2	Conditions and Example Items for Experiment 4 (OPP)	99
Table 7.3	Omnibus RNP Analyses for Experiment 3	101
Table 7.4	Separate RNP Analysis for Experiment 3: Null Pronoun vs. Repeated-name Non-topic- <i>ga</i>	102
Table 7.5	Separate RNP Analysis for Experiment 3: Null Pronoun vs. Repeated-name Topic- <i>wa</i>	103
Table 7.6	Simple Topic vs. Non-topic Antecedent Effects for Each Anaphor Form in Experiment 3	104
Table 7.7	Omnibus OPP Analyses for Experiment 4	106
Table 7.8	Separate OPP Analysis for Experiment 4: Null Pronoun vs. Overt-pronoun Non-topic- <i>ga</i>	107
Table 7.9	Separate OPP Analysis for Experiment 4: Null Pronoun vs. Overt-pronoun Topic- <i>wa</i>	108
Table 7.10	Simple Topic vs. Non-topic Antecedent Effects for Each Anaphor Form in Experiment 4	109
Table 8.1	Verbs in Items of Experiment 5 and 6	117
Table 8.2	Conditions and Example Items for Experiment 5 (RNP)	119
Table 8.3	Conditions and Example Items for Experiment 6 (OPP)	120
Table 8.4	Omnibus RNP Analyses for Experiment 5	123
Table 8.5	Separate RNP Analysis for Experiment 5: Null Pronoun vs. Repeated-name Non-topic- <i>ga</i>	124
Table 8.6	Separate RNP Analysis for Experiment 5: Null Pronoun vs. Repeated-name Topic- <i>wa</i>	125

Table 8.7	Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 5	126
Table 8.8	Omnibus OPP Analyses for Experiment 6	128
Table 8.9	Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Non-topic- <i>ga</i>	129
Table 8.10	Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-Pronoun Topic- <i>wa</i>	130
Table 8.11	Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 6	131
Table 8.12	Omnibus OPP Analyses for Experiment 6 with <i>Ageru/Kureru</i> ‘give’ & ‘do a favor of –ing’	134
Table 8.13	Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Non-topic- <i>ga</i> with <i>Ageru/Kureru</i> ‘give’ & ‘do a favor of –ing’	135
Table 8.14	Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Non-topic- <i>wa</i> with <i>Ageru/Kureru</i> ‘give’ & ‘do a favor of –ing’	136
Table 8.15	Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 6 with <i>Ageru/Kureru</i> ‘give’ & ‘do a favor of –ing’	137

List of Figures

Figure 2.1.	Referents allowed to be referred to with <i>kare</i> ‘he’ and <i>kanojo</i> ‘she’	27
Figure 6.1.	Reading times of the second sentences with anaphors in Experiment 1	82
Figure 6.2.	Reading times of the second sentences with anaphors in Experiment 2	87
Figure 7.1.	Reading times of the second sentences with anaphors in Experiment 3	100
Figure 7.2.	Reading times of the second sentences with anaphors in Experiment 4	105
Figure 8.1.	Reading times of the second sentences with anaphors in Experiment 5	122
Figure 8.2.	Reading times of the second sentences with anaphors in Experiment 6	127
Figure 8.3.	Reading times of the second sentences with anaphors in Experiment 6 with items <i>ageru/yaru</i> & <i>kureru</i>	133

Chapter 1. Introduction

The primary objective of this study is to investigate the effects of the Repeated Name Penalty (RNP) and the Overt Pronoun Penalty (OPP) in Japanese. The RNP was first reported by Gordon, Grosz and Gilliom (1993), who investigated the processing of personal pronouns and repeated names when these anaphors refer to person referents. They observed that when anaphoric reference appears in the subject position of a sentence, sentences with repetitive full referential noun-phrase (NP) anaphors (i.e., repeated names of the referent person) were read slower than pronouns. This slowed reading time is not attributed to the shorter length of pronouns relative to repeated names because it was only observed with subject anaphors referring to antecedents that were either the grammatical subject (e.g., *John gave Mary a pet hamster. John/He didn't have a cage to put it in.*) or the first-mentioned surface-initial NP (e.g., *In John's opinion, Mary should not have given him a pet. John/He didn't have a cage to put it in.*). In contrast, when their antecedents were the grammatical object, the form of an anaphor (i.e., pronoun, repeated name) did not affect the speed with which the sentence was read. This interaction between anaphors and antecedents is the indication of the RNP. On the other hand, studies of a number of null subject languages such as Spanish (Gelormini-Lezama & Almor, 2011), Italian, and Brazilian Portuguese (Almor, de Carvalho Maia, Cunha Lima, Gelormini-Lezama & Vernice, 2013) found that overt-pronoun subject sentences (as well as repeated-name subject sentences) were read slower than null-pronoun subject sentences when referring to subject antecedents (e.g., *John gave Mary a pet hamster.*

John/He/Ø didn't have a cage to put it in.). Gelormini-Lezama and Almor (2011) named this slowed reading of overt pronouns relative to null pronouns the Overt Pronoun Penalty (OPP). The current study investigated if we detect these processing penalties with different anaphors referring to different antecedents in Japanese.¹

In addition to the anaphor forms investigated in the previous studies above (i.e., repeated names, overt pronouns and null pronouns), the present study also examined a discourse-based NP category, the topic, as an anaphor. Previous studies examining the RNP and OPP observed slower readings of repeated-name subjects relative to overt-pronoun subjects and slower readings of overt-pronoun subjects relative to null-pronoun subjects. However, it still remains unclear whether these effects reflect a grammatical constraint associated with anaphors' grammatical subject-hood or an information structure constraint associated with anaphors' topic-hood. This lack of clarity likely reflects the fact that the subject and topic were not overtly differentiated in the languages examined in the previous studies. In Japanese, the topic is expressed either in the form of a null pronoun or as an NP with the topic postposition, *wa*.² Both forms are distinct from non-topic NPs, which are not null or not marked by *wa*. By overtly differentiating the topic NP and non-topic NP, the present study measured the reading times of sentences with the topic or non-topic anaphors that were repeated names, overt pronouns and null pronouns. In sum, five variations of anaphors were used in this study: repeated-name

¹ This dissertation uses the terms, *anaphor* as an element that refers to an entity that was introduced before, and *antecedent* as an element that was introduced before.

² In some linguistic studies, *postposition* is differentiated from *marker*. This dissertation does not make this distinction.

topic, repeated-name non-topic, overt-pronoun topic, overt-pronoun non-topic, and null-pronoun topic.³ Following Gelormini-Lezama and Almor (2011), for the examination of RNP, reading times of repeated-name-topic/repeated-name-non-topic sentences and null-pronoun sentences were compared; for the examination of OPP, reading times of overt-pronoun-topic/overt-pronoun-non-topic sentences and null-pronoun sentences were compared. I looked for the interactions of these anaphors above and different types of antecedents as the indications of RNP and OPP (baseline: null pronouns for anaphors, salient antecedents for antecedents).

The reading times of anaphors are also affected by the salience of antecedents. In the previous studies, slower reading occurred when the antecedents were the grammatical subjects or first-mentioned NP in the sentence. In other words, the relative slower reading of certain types of anaphors (e.g., English repeated-name anaphors, Spanish overt pronouns) is observed when the antecedents are highly salient. Salience refers to the

³ In some syntactic perspectives, the structure of null topic sentences involve an empty operator in the complementizer phrase (CP) spec position. For example, the sentence *Moo tabeta* ‘already ate’ with an understood topic is, [CP Op_{[TOP]i} [IP pro_i moo [VP tabe]ta]. However, this dissertation does not concern the syntactic position of the topic as Newmeyer (1999) and Lambrecht (1994) state that it is impossible for grammars to directly mirror any discourse function. Therefore, this dissertation treats null-pronoun topic subjects and overt NP-*wa* topic subjects only as the discourse topic and grammatical subject (i.e., [IP NP-*wa*/pro [VP V]]). Also, non-topic-subject anaphors with *ga* and topic-subject anaphors with *wa* are both dealt with as the subject, and the difference between them are only postpositions.

degree of activation in readers' memory; an entity being salient means that the entity is highly activated in readers' working memory. The salience of an antecedent is determined by its syntactic category and by the sentence's word order (e.g., subjects are more salient than non-subjects; first-mentioned NPs are more salient than the other NPs). The change in reading speeds based on the salience of the antecedent reflects the different levels of ease of retrieval for the antecedent. Ariel (1990) calls the level of the ease of antecedents' retrieval *accessibility*. Grosz, Joshi and Weinstein's (1995) Centering Theory suggests an accessibility-ranking of antecedents based on the antecedents' grammatical categories: Subject > Object > Others, from the most accessible to the least accessible. For example, in the sentence shown below, the subject *John* is more accessible than the object, *Chris*, i.e., *John* is easier to retrieve than *Chris*.

(1) John slapped Chris. He...

While Grosz, et al.'s ranking is based on the grammatical categories of antecedent-NPs, Ariel (1990) and Gundel, Hedberg and Zacharski (1993) maintain that antecedents' accessibilities and salientces are enhanced by the topic-hood of the antecedents (i.e., Topic > Non-Topic). For example, in the following discourses, both the topic antecedent and non-topic antecedent are grammatical subjects below. However for the anaphor *He*, the topic antecedent *he* in (2a) might be more accessible than the non-topic antecedent *Peter* in (2b).

(2) a. As for Peter, he slapped Mary. He....

b. Peter slapped Mary. He....

In addition, Walker, Iida and Cote (1994) suggest that the accessibility is also enhanced by empathy (i.e., Empathy locus > Non-empathy locus). In other words, an

entity that the speaker is empathizing with is more accessible than the other entities. For instance, in (3a) below, the speaker uses the term [*Taro's*] *daughter* to describe *Hanako*, so the speaker is empathizing more with *Taro* than *Hanako* (i.e., *Taro* is the speaker's empathy locus). In contrast, in (3b), the speaker states *Hanako's father* to describe *Taro*, so the speaker is empathizing more with *Hanako* than *Taro* (*Hanako* is the speaker's empathy locus).

- (3) a. Taro hugged his daughter Hanako. He....
 b. Hanako's father Taro hugged her. He....

Taro is the grammatical subject in the first sentence of both discourses, so for the anaphor *He*, *Taro* should be equally accessible in (3a) and (3b). However, the accessibility of *Taro* could be lower in (3b) relative to (3a).

The current study tested the saliences and accessibilities of these different antecedents: subject vs. object, topic vs. non-topic, and empathy locus vs. non-empathy locus. These antecedents' accessibilities are examined along with the five different anaphors listed earlier (i.e. repeated-name topic, repeated-name non-topic, overt-pronoun topic, overt-pronoun non-topic, and null-pronoun topic).

The present study investigates processing of anaphors that refer to a person referent, using the different types of anaphors and antecedents listed above in order to investigate the RNP and OPP in Japanese.⁴ This study conducted a series of self-paced

⁴ Following Gordon et al. (1993) and Gelormini-Lezama and Almor (2011), this study focuses on the processing of anaphors that are either repeated names, overt pronouns or null pronouns. This study does not deal with common nouns such as the ones that accompany definite articles (e.g., "Dr. Smith... The professor...").

reading experiments, following Gordon, Grosz and Gilliom (1993) and Gelormini-Lezama and Almor (2011). To my knowledge, there are relatively few experimental studies on the online processing of different anaphors and antecedents in Japanese or on the influence of topic and empathy. This study aims to fill this gap in the literature, contributing to a cross-language understanding of reference.

This dissertation is organized as follows. Chapter 2 discusses anaphor forms, particularly repeated names, overt pronouns and null pronouns, and reviews earlier studies that tested different processing patterns with these anaphors. The chapter also includes a detailed review of pronouns in Japanese. Chapter 3 describes postpositions in Japanese. Because postpositions mark anaphors, they could affect anaphor processing. Chapter 4 discusses antecedents. The chapter focuses on three different contrasts of antecedent types, which are mentioned above: grammatical subjects vs. non-subjects, discourse-topic vs. non-topic, and empathy locus vs. non-empathy locus. Chapter 5 describes a broad overview of the six experiments for the current study. Chapter 6 describes Experiments 1 and 2, which tested the RNP (Experiment 1) and the OPP (Experiment 2) with subject antecedents vs. non-subject antecedents. Chapter 7 describes Experiments 3 and 4, which tested the RNP (Experiment 3) and the OPP (Experiment 4) with topic antecedents vs. non-topic-antecedents. Chapter 8 describes Experiments 5 and 6, which tested the RNP (Experiment 5) and the OPP (Experiment 6) with empathy-locus antecedents vs. non-empathy-locus antecedents. Chapter 9 provides the general discussion on all the experiments and their outcomes, and Chapter 10 concludes this dissertation.

Chapter 2. Anaphors

Anaphors, the Repeated Name Penalty, and the Overt Pronoun Penalty

It has been widely argued that different anaphor forms such as repeated names, overt pronouns and null pronouns differently affect reference production and resolution. A number of early studies on English found the preference for using pronouns in substitution for repetitive full NPs when referring to antecedents. For example, Sanford, Moar and Garrod (1988) suggested that, while pronouns refer to antecedents, full NPs generally introduce new entities into discourse. In addition, Cloitre and Bever's (1988) study found that readers process the reference faster when an anaphor is a pronoun rather than when it is a full NP.

On the other hand, Gernsbacher's (1990) research on repeated names and personal pronouns claimed that the more that the information denoted by an anaphor overlaps with the information of its antecedent, the easier it is to realize their referential relationship. Thus, she suggested that informative anaphors such as repeated names have an advantage over less informative ones such as pronouns because a repeated-name anaphor includes information that fully matches the information of its antecedent. According to this view, repeated-name anaphors activate the antecedents most effectively and should be processed faster than the anaphors in less informative forms (e.g. pronouns).

However, Gordon, Grosz and Gilliom (1993) showed the processing disadvantage of repeated-name anaphors over pronouns in English. They observed that, when an anaphoric reference appears in the subject position of a sentence, sentences with

repeated-name anaphors are read slower than comparable sentences with pronoun anaphors (e.g., *John gave Mary a pet hamster. John/He didn't have a cage to put it in.*). Gordon et al. named this slower reading of repeated-name anaphors the Repeated Name Penalty (RNP). They found that slowed reading of repeated names only occurs for subject anaphors that refer to antecedents which are either positioned high in the syntactic tree, namely as the subject (e.g., *John gave Mary a pet hamster. John/He didn't have a cage to put it in.*) or as the first-mentioned surface-initial NP (e.g., *In John's opinion, Mary should not have given him a pet. John/He didn't have a cage to put it in.*). A summary of Gordon et al.'s findings is shown in Table 2.1 below.

Table 2.1

The Repeated Name Penalty (Gordon et al., 1993)

Experiments	Antecedent	Anaphor	Results
Experiment 1	Subject	Repeated name (Subject)	Slower reading
		Pronoun (Subject)	
	Object	Repeated name (Object)	
		Pronoun (Object)	
Experiment 2	Subject	Repeated name (Surface-initial Non-subject)	
		Pronoun (Surface-initial Non-subject)	
Experiment 3	Subject	Repeated name (Non-surface-initial Subject)	Slower reading
		Pronoun (Non-surface-initial Subject)	
	Object	Repeated name (Surface-initial Non-subject)	
		Pronoun (Surface-initial Non-subject)	

Experiment 4	Subject	Repeated name (Subject)	Slower reading
		Pronoun (Subject)	
	Object	Repeated name (Subject)	
		Pronoun (Subject)	
Experiment 5	Non-surface-initial Subject	Repeated name (Subject)	Slower reading
		Pronoun (Subject)	
	Surface-initial Non-subject	Repeated name (Subject)	Slower reading
		Pronoun (Subject)	

Gordon et al. (1993) used two-sentence discourses in this study above, but Gordon, Hendrick, Ledoux and Yang (1999) found that the same processing penalty in a single sentence that includes two clauses (e.g., *John went to the store so that John/he could buy candy.*); an antecedent in the first clause and an anaphor in the second clause.⁵

Although the studies described above investigated the differential processing of repeated names and pronoun anaphors, anaphors' forms are not limited to repeated names and overt pronouns in many other languages. For example, Spanish and Italian utilize null pronouns for the subject in addition to overt pronouns. Several studies on the null and overt pronouns in Romance languages seem to show equivalences between English overt pronouns and Spanish/Italian null pronouns and between English repeated names and Spanish/Italian overt pronouns. In other words, English overt-pronoun subjects and Spanish/Italian null-pronoun subjects are preferred to English repeated-name subjects and Spanish/Italian overt-pronoun subjects, respectively, when referring to salient antecedents.

⁵ Experiments in this dissertation solely used two-sentence discourses.

For example, Alonso-Ovalle, Fernández-Solera, Frazier and Clinton (2002) examined differences in the use of null pronouns and overt pronouns in Spanish. Based upon the findings of their written questionnaire study, native Spanish speakers used null pronouns to refer to subject antecedents 73.2% of the time but used overt pronouns only 50.2% of the time to refer to subject antecedents. Also, in the acceptability judgment tests from the same study, sentences using null-pronoun anaphors to refer to subject antecedents were seen as significantly more acceptable than those using overt pronouns. In addition, Sorace and Filiaci's (2006) sentence-interpretation experiment in Italian used null and overt pronouns referring to ambiguous antecedents. The results showed that native Italian-speaking participants tended to interpret the null-pronoun subjects referring to the subject antecedents, and tended to interpret the overt-pronoun subjects referring to the object antecedents. Belletti, Bennati and Sorace (2007) state that, "typically, an overt pronominal subject of an embedded clause does not refer to the preverbal lexical subject of a superordinate clause..., whereas a null subject does" (p. 660).

Gelormini-Lezama and Almor (2011) conducted a self-paced reading study in Spanish, following Gordon, Grosz and Gilliom (1993). They found that sentences with repeated-name or overt-pronoun subject anaphors were read slower than sentences with null-pronoun subject anaphors when the antecedents were in the subject position of the previous sentence (e.g., *Juan se encontró con María. Juan/El/Ø la vio triste.* 'Juan met with Maria. Juan/He/Ø found her sad.'). They argued that the slower reading of the repeated names reflects an RNP in Spanish whereas the slower reading of overt pronouns, which they called the Overt Pronoun Penalty (OPP), is specific to null subject languages. While Gelormini-Lezama and Almor observed the OPP with subject antecedents, the

effect was reversed when the antecedents were the object; sentences with null-pronoun anaphors were read slower than sentences with overt pronouns and repeated names. Their findings are summarized below.

Table 2.2

The Overt Pronoun Penalty (Gelormini-Lezama & Almor, 2011)

Experiment	Antecedent	Anaphor	Results
Experiment 1	Subject	Repeated Name (Subject)	Slower reading
		Overt Pronoun (Subject)	Slower reading
		Null Pronoun (Subject)	
	Object	Repeated Name (Subject)	
		Overt Pronoun (Subject)	
		Null Pronoun (Subject)	Slower reading

Gordon, Grosz and Gilliom (1993) originally explained the RNP in the framework of Centering Theory (Grosz, Joshi & Weinstein, 1995), arguing that this effect violates a principle mandating the use of pronouns to refer to the most prominent entity in the discourse. Gordon and Hendrick (1998) later modified this explanation into Discourse Prominence Theory (DPT).⁶ They argued that repeated names are initially interpreted

⁶ Gordon and Hendrick (1997) define antecedents' prominence as, if X c-commands Y, and if Y does not c-command X, X is more prominent than Y. Thus, grammatical

non-anaphorically such that the reader of a repeated-name anaphor first establishes a new representation of the entity referred to by the repeated name and only then searches for a possible matching antecedent, starting with the least prominent antecedent and proceeding in increasing order of antecedent prominence. In contrast, pronouns are immediately interpreted as anaphors that need antecedents, which are searched for in decreasing order of antecedent prominence, starting with the most prominent antecedent.⁷ Thus, for prominent antecedents, pronouns are processed faster than repeated-name anaphors. However, the DPT cannot explain the different distributions of null and overt pronouns in Spanish and Italian, which are detected by Frazier and Clinton (2002), Sorace and Filiaci (2006) and Gelormini-Lezama and Almor (2011). Specifically, when antecedents are prominent ones such as grammatical subjects, the use of overt pronouns is inappropriate. This does not agree with the DPT because overt pronouns should be immediately interpreted as anaphors as long as they are pronouns, and thus null and overt pronouns should behave similarly.

Almor (1999) offered an alternative explanation for the advantage of using pronouns over repeated names when referring to prominent (or salient) antecedents. Almor's explanation is based on the Grice's (1975) Maxim of Quantity, which insists on the pragmatic functionality of an utterance: "Make your contribution as informative as required" (Grice, 1975, p. 45). According to Almor's Informational Load Hypothesis

subjects are more prominent than objects since subjects c-command objects, and objects do not c-command subjects.

⁷ This dissertation considers that the words *salience* and *prominence* roughly mean the same.

(ILH), the amount of semantic information that an anaphor carries must be just enough to identify the referent and, in some cases, add new information to the discourse. When the subject is the intended antecedent, because the subject is the default antecedent, an anaphor referring to the subject does not have to carry much semantic information that repeated-name anaphors carry. Therefore, using repeated-name anaphors for subject antecedents provides too much semantic information that serves no discourse function, slowing reading. On the other hand, the RNP does not occur when the antecedent is not the subject, because the semantic information carried by a repeated-name anaphor then serves a discourse function in helping readers avoid the default antecedent (i.e., subject) and identify a non-subject NP as the antecedent.

The same explanation can be applied to Gelormini-Lezama and Almor's (2011) findings of both the RNP and OPP in Spanish. Gelormini-Lezama and Almor found that Spanish overt pronouns and repeated names lead to equally slower readings than null pronouns. These results suggest that both overt pronouns and repeated names in Spanish carry more semantic information than is necessary in order to identify subject antecedents. The information that Spanish overt pronouns carry is redundant because their gender and number features overlap with the information that Spanish verbal morphology provides. The information carried by rich verbal morphology and by overt pronouns accumulatively could reach a level where they penalize readers' processing. This argument is supported by the fact that the OPP is not observed in Chinese, which does not employ verbal morphology (Yang, Gordon, Hendrick & Wu, 1999). On the other hand, when antecedents were grammatical objects, Spanish sentences with overt pronouns and repeated names were read faster than those with null pronouns. This result

indicates that readers needed informative anaphors to avoid the default antecedent (i.e., subject) and identify the object antecedent; overt pronouns and repeated names were satisfactory informative but null pronouns were not.

Using overt-pronoun anaphors in Spanish referring to salient subject antecedents will be accepted in a given context where the anaphor is focused (Larson & Lujàn, 1989).

- (4) a. Cuando Juan trabaja, *él/Ø no bebe.
 when John work, he/Ø not drink.
 ‘When John works, he doesn’t drink.’
- b. (John allows other people to drink when they work. However,)
 Cuando Juan trabaja, él/Ø no bebe.
 when John work, he/Ø not drink.
 ‘when John works, he doesn’t drink.’

(Larson & Lujàn, 1989, p. 13)

In (4a) above, the use of an overt pronoun is avoided and a null pronoun is preferred instead, as Gelormini-Lezama and Almor’s (2011) OPP experiments empirically revealed. However, in (4b) above, the overt pronoun is read natural in the context that *John* is focused in the contrast with the other workers. This usage of Spanish overt pronouns is accepted because it contributes to the identification of the antecedents among competing entities. In this regard, Spanish overt pronouns parallel English stressed pronouns, as shown below (Larson & Lujàn, 1989). In (5a) below, the overt pronoun *he* cannot be stressed, but it can be in (5b) where John is focused among multiple people.

- (5) a. When John works, *HE/he doesn’t drink.
- b. (John allows other people to drink when they work, but)
 when John works, HE/he doesn’t drink.

(Larson & Lujàn, 1989, pp. 4-5)

Likewise, both English and Spanish should allow the use of repeated-name anaphors when they focus one of multiple antecedents. Studies regarding the RNP in English (Gordon, Hendrick, Ledoux & Yang, 1999; Swaab, Camblin & Gordon, 2004; Nair & Almor, 2009) illustrate that reading speed does not slow when a repeated-name anaphor focuses one of the equally salient antecedents in a conjoined phrase (e.g., *John and Mary went to the store. John wished to buy some candy.*). Also, the original OPP study by Gelormini-Lezama and Almor (2011) shows that neither the RNP nor OPP occurs when Spanish repeated-name and overt-pronoun anaphors are the discourse focus. In their Experiment 2, they explicitly gave a focus-hood on anaphors by utilizing the it-cleft structure (e.g., *John met with Maria. It was John/he/Ø who found her sad.*). The results showed that the sentences with repeated names were read the fastest, sentences with overt pronouns were read the second fastest, and those with null pronouns were read the slowest. These outcomes demonstrate that, when an anaphor is a focused entity, the anaphor is preferred to be in an overt form (i.e., English and Spanish repeated names and Spanish overt pronouns), contrary to the RNP and OPP.

Overt Pronoun in Japanese

Japanese utilizes both overt pronouns and null pronouns. Gundel, Hedberg and Zacharski's (1993) Givenness Hierarchy shows the equivalent anaphors' forms across five languages including English, Spanish and Japanese, as shown below. According to the Givenness Hierarchy, English stressed pronouns, Spanish overt pronouns and Japanese overt pronouns are equivalent, which refer to antecedents activated in working memory. However, a contradictory fact is that, unlike English and Spanish speakers, native Japanese speakers rarely use the 2nd and 3rd person overt pronouns; instead they

tend to use null pronouns and full NPs including referents' names, their titles (e.g., "teacher") and kinship terms (e.g., "mother") (Shibatani, 1990).

Table 2.3

The Givenness Hierarchy: Equivalent Anaphor Forms in English, Spanish and Japanese (Gundel et al., 1993, p. 284)

<u>Language</u>	<u>Anaphor</u>		
	for more salient antecedent < ----- > for less salient antecedents		
English	(unstressed) he/she	(stressed) HE/SHE	the NP
Spanish	Ø	él/ella	el NP
Japanese	Ø	kare/kanojo	NP

Clancy's (1980) corpus study finds that, when referring to a 3rd person, Japanese speakers' narratives employed null pronouns (73.2%) and full NPs (26.8%) but never used overt pronouns. Nagano's (2012) corpus study shows that, when referring to a 3rd person, only 0.5% of predicates used an overt pronoun in her oral interview with native Japanese speakers. Miyaji's (1964) corpus study with Japanese magazines finds that 90.15% of references to the 1st person were made with overt pronouns, but 28.47% of references to the 2nd person and only 0.15% of references to the 3rd person were made by overt pronouns. Almost all of the other references without overt pronouns were made by referents' names, titles and kinship terms. Suzuki (1973/1978) argues, "there is a definite

tendency to avoid [the overt-pronoun] use as often as possible and to carry on conversation using some other words to designate speaker and address” (p. 92).

It is of importance to note that the class of expressions referred to as “Japanese overt pronouns” is syntactically, semantically, and pragmatically different from overt pronouns in many other languages including English and Spanish, as Clancy (1980) and Hirose (2000) maintain that Japanese does not have a system of overt pronouns corresponding directly to those of English. Further, in observing the distribution of Japanese overt pronouns, they might not be “pronouns” in the traditional sense. One of the unique properties of Japanese overt pronouns that are distinct from other languages’ pronouns is that they can be modified by a clause or an adjective, just like regular nouns. In other words, it is grammatical in Japanese to say, “he, whom I saw yesterday” or “beautiful she”. Also, Japanese overt pronouns can be modified by demonstratives; it is grammatical in Japanese to say, “this I” or “that he”. Another example of uniqueness of Japanese overt pronouns is their length. In general, pronouns should be shorter than full NPs for the purpose of communicative efficacy, but most Japanese overt pronouns such as *anata* ‘you’ and *kanojo* ‘she’ are neither orthographically nor phonologically shorter than a typical regular noun. Considering that “there is strong evidence that [...] word forms contribute to processing difficulty” (Jaeger & Tily, 2010, p. 326), Japanese overt pronouns would not contribute to efficient communication. In addition, it is known that Japanese 1st and 2nd person overt pronouns have a large number of forms, and most of the forms of Japanese overt pronouns derive from regular nouns (e.g., *watakusi* ‘I’ originally meant ‘private thing’; *kimi* ‘you’ meant ‘emperor’) or from deictic terms (e.g., *anata* ‘you’ and *kare* ‘he’ formerly meant ‘that place in the distance’ or ‘that person in the distance’)

(Miyaji, 1964; Shibatani, 1990). Moreover, the use of each form is subject to socio-cultural factors to a greater extent compared with other languages (Hinds, 1975; Kanaya, 2002; Obana, 2003a, b; Ono & Thompson, 2004; Palmroos, 2010).

1st person Overt Pronoun

Japanese 1st person pronouns (i.e., ‘I’ in English) include *watasi*, *atasi*, *watakusi*, *atakusi*, *watai*, *atai*, *wate*, *ate*, *boku*, *ore*, *ora*, *oira*, *wai*, *wasi*, *assi*, *oidon*, *uchi*, *zibun*, *ware*, *soregasi*, *sessha*, *temae*, *onore* and *shosei*. These variations are correlated with socio-cultural factors such as the gender, age, and relative social standing of the speaker and the addressee, as well as with other discourse-related and language-related variables such as speech level (i.e. formality-familiarity), and dialectal differences. For example, *watasi* and *watakusi*, which are perhaps the most general 1st person pronouns, are the only forms normally used in formal situations such as business meetings, ceremonial speeches, or employment interviews.⁸ If one does not call himself/herself *watasi* or *watakusi* in an interview, and instead uses a cruder, more vernacular form of the 1st person pronoun, such as *ore*, he/she would most likely not be hired, at least in the white-collar job market. The 1st person forms *atasi*, *atakusi*, *atai* and *ate* are all variants of *watasi*, *watakusi*, *watai* and *wate* with the deletion of the word-initial “w” (i.e., syncope). Those forms without “w” are used in more informal situations than those with “w”. Further, while the forms with “w” are used by both male and female speakers, those without “w” are used only by females. If a male speaker calls himself *atasi*, he will sound

⁸ Although *watasi* and *watakusi* are formal versions of the 1st person pronoun, there is a difference of the degree of formality between these two, i.e., *watakusi* is more formal than *watasi*.

to some extent as though he was presenting himself as a female (e.g., he might be transsexual). On the other hand, *boku*, *ore*, *ora*, *oira*, *wai*, *wasi*, *assi* and *shosei* are used only by males. Among these words, *boku* and *ore* are the most common, but *ore* sounds rougher, more masculine and more informal than *boku*. Shibatani (1990) shows the differentiation of the uses of the 1st person pronouns based on gender and formality in the table below.

Table 2.4

Distinction in Pronouns (Shibatani, 1990, p. 371)

	Formal < ----- > Informal			
1 st person				
Male speaker	watakusi	watasi	boku	ore
Female speaker	watakusi	watasi		atasi

However, gender and formality are not the only factors that determine the use of Japanese overt pronouns. Age is another factor that comes to bear in determining the use of the 1st person overt pronoun forms. For example, as *wasi* is typically used by older males at or around retirement age, and *boku* is mainly used by younger males. The 1st person form *boku* was initially a Chinese loanword meaning ‘slave/servant’, and first developed as a student slang expression meaning ‘your servant = I/me’ (Martin, 1975, p. 1076). Since

then, *boku* has come to express a speaker's deference to his addressee, communicating the sense that he is young, and less experienced than his addressee. If a highly experienced or knowledgeable person with a high social status calls himself *boku*, that means that he is bringing himself to a lower status in order to be humble or to fill the gap of social statuses between himself and the addressee (i.e., the speaker is attempting to speak in a frank and friendly manner). The forms *ora* and *oira* ordinarily communicate that the speaker is a rustic individual (e.g., is from rural or countryside regions and likely from Northern Japan).⁹ The forms *uchi* (which is also used as a regular noun that means 'house/home') and *wai* are dialectic terms of Western Japan, with *uchi* typically used by females, and *wai* by males. The form *oidon* is also a dialect-specific 1st person pronoun used by males in the southernmost region of Japan (i.e., Kyushu region). It is also a form that is felt to be out of fashion, and its use has fallen off drastically in recent years. Finally, the term *zibun* is often categorized by Japanese linguists as a reflexive pronoun and is most commonly used as a 1st person pronoun by the people who have participated in physically strict training regimens, such as in the Self-Defense Forces or a traditional athletic team.¹⁰ *Zibun* as a 1st person pronoun became popular among military men before

⁹ A well-known example is found in a popular cartoon, *Dragon Ball*. Goku, the protagonist, who grew up in a mountain, calls himself *ora*. In contrast, his son, Gohan, calls himself *boku*, which reflects the influence from his mother who is obsessed with his education.

¹⁰ The Self-Defense Forces are the military forces of Japan. Its activity is limited to reactive defenses against attacks from outsiders and is prohibited from the participation in a battle outside Japan.

Japan's defeat in World War II. The use of *zibun* as the 1st person pronoun allows a speaker to present himself as being traditional, and possessing a strong masculine character. The terms *soregasi*, *temae* and *onore* were used a few hundred years ago, by the people in high social classes (e.g. such as members of the samurai class), but very few people now use these terms. The term *shosei* is extremely formal and exclusively used in formal writing.

A Japanese speaker uses these different 1st person pronouns depending on different situations. A boy who was born in countryside might call himself *ora*. If he moves to a big city, he could change *ora* to *boku* and/or *ore* in an attempt to fix or hide his rural background. He might use *ore* when talking with his classmates and *boku* when talking with his school principal. If he becomes a businessman, he would have to use *watasi/watakusi* (but he could still keep using *ore* when talking to his close friends or family members). When he gets old and is retired, he might start calling himself *wasi*. If he writes his autobiography before he dies, he might refer to himself in it using the pronoun *shosei*. Summarizing, Japanese “1st person overt pronouns” are much more than mere pronouns. They express a speaker's gender, age, background, his relationship and standing respective of the addressee, the level of formality and politeness particular to the discourse context, and sometimes the mode of expression (i.e., written vs. spoken).

2nd person Overt Pronoun

Japanese 2nd person pronouns (i.e. ‘you’ in English) also have a large number of forms. They include *anata*, *anta*, *kimi*, *omae*, *omee*, *temee*, *kisama*, *onore*, *ondore*, *odore*, *ware*, *zibun*, *boku*, *onusi*, *nanji*, *unu*, *kikei*, *kiden* and *otaku*. Similar to the 1st person pronouns, their use is determined by socio-cultural and discourse factors. However, 2nd

person pronouns are much more circumscribed in their use. That is, although these forms can express different degrees of formality and politeness toward the addressee, none of the 2nd person pronouns are simply not to be used when the addressee is older or of a higher status than the speaker (Shibatani, 1990). For example, Tanaka (1993) recounts a situation wherein a Chinese worker, an employee at a Japanese company founded in China, called his Japanese boss *anata*, and the worker was fired on account of it. This is because it is considered highly disrespectful to address a social superior with the use of an overt pronoun. Also, the 2nd person overt pronoun is not to be used in referring to an addressee who is older than the speaker. For instance, Hasumi (1986) describes a situation in which he was shocked by his Japanese-French bilingual child addressing him as *anata* in Japanese. When addressing a social superior or someone who is older, Japanese politeness conventions dictate that one should use their actual names (e.g., an individual named *Honda Jiro* would be addressed as *Honda-san* ‘Mr. Honda’ by someone who was not a close friend or acquaintance, and as *Jiro-san* by someone who was so). Alternatively, one may address a person using their title (plus the respect ending *-san*). Thus, one’s *otou* ‘father’ would be *otou-san* and one’s *sensei* ‘teacher’ might be *Sensei*, *Honda-sensei*, or *Jiro-sensei*.

The most common 2nd person pronouns are *anata*, *anta*, *kimi* and *omae*.¹¹ Among these, *anata* and *kimi* are used in formal situations, and *anta* and *omae* are informal. For instance, as an example of a formal situation, a professor can call his/her student *anata* or

¹¹ As mentioned earlier, *anata* was originally a deictic term. *Anata* meant ‘a person in distance/a distant place’; *konata* meant ‘the speaker/a place near the speaker’; *sonata* meant ‘the listener/a place near the listener’.

kimi; if a professor calls a student *anta* or *omae*, the professor might appear to be either overly friendly with or perhaps even rude to the student. As an example of an appropriately informal situation, a parent is entitled to call his/her child *anta* or *omae*; if a parent calls his/her child *anata* or *kimi*, their relationship would seem overly formal.

Japanese also has a number of 2nd person pronoun forms that have clearly negative connotations and are used as insults or terms of disrespect. Included among these are *kisama*, *temee*, *onore*, *ondore*, *odore* and *ware*. These forms are typically used when a speaker is angrily yelling at or arrogantly insulting an addressee. They can be thought of as equivalent to ‘you, bastard!’ in English. Note that these forms are derived from two older 1st person pronoun forms, *temae* and *onore*. When used as 1st person pronouns, they would communicate a reserved and humble disposition on the part of the speaker. Thus, when used as 2nd person pronoun, their effect is the opposite. In this case they communicate that the speaker deems the addressee as needing to be humble – which is an insult, from a politeness perspective. In other words, by lowering the addressee’s status rather than his own the speaker is seen to be looking down upon the addressee.

Interestingly, the meaning of *kisama* has changed over time; it originally expressed a measure politeness toward an addressee, but is now used to insult. While *zibun* is used (as we have seen) as a 1st person pronoun, it is also used as the 2nd person pronoun by speakers of the Kansai regional dialect. Similarly, *boku* is a 1st person pronoun for young males, but can be used as a 2nd person pronoun when talking to a small male child. *Onusi*, *nanji* and *unu* are classical terms from pre-modern era, and are rarely used now. *Kikei* and *kiden* are mostly used in formal writing, and express a speaker’s (or writer’s) politeness toward the referent. *Otaku* is different from the others in that, while most

Japanese 2nd person pronouns' referents have to be persons, the referent of *otaku* can be also an organization or institute (similar to English 'you'). For example, one can say *otaku* in referring to a company when talking with the telephone operator of the company, as "I thought you (*otaku*) accept the payment by a credit card".

3rd person Overt Pronoun

Unlike the 1st and 2nd person pronouns, Japanese 3rd person pronouns have only two basic forms, namely *kare* 'he' and *kanojo* 'she'. In this regard, Japanese 3rd person overt pronouns are more comparable to those of English than are their 1st and the 2nd person pronouns (Ishiguro, 2013). The reason that there are so few variant of the 3rd person pronominal form is that Japanese did not traditionally ever have anything that might be considered an overt 3rd person pronoun, and that they (these pronouns) were originally coined for the purpose of translating texts written in Western languages (Shibatani, 1991; Martin, 1975). In the mid-19th century, Japan was forced to abandon its isolationist policy and began thenceforth to develop itself as an industrial and military power. In this context, Japan began importing ideas (and texts) from the West, and translators found the need to coin gendered pronouns that are equivalent to 'he' and 'she' in Western languages.

The 3rd person overt pronouns *kare* 'he' and *kanojo* 'she' are derived from a noun, *kare/ka*, which meant 'a distant location from a speaker'. When *kare/ka* was used to refer to a person-entity, it indicated a person who was physically or psychologically distant from the speaker, as a sign of politeness to the person. The word, *kare* 'distant location/person in distance', was adapted as the counterpart of English 'he', and *kanojo* 'she' was derived from adaptation of *ka* 'distant location/person in distance' plus the

genitive-postposition *no* and *onna/jo* ‘woman’, as shown below (Obana, 2003b; Ishiguro, 2013).¹²

- (6) kare/ka
distant location/person in distance
- kare
distant location/person in distance
'he'
- ka
distant location/person in distance
'she'
- no onna/jo
GEN woman

The construction *ka-no NP* is still found in contemporary Japanese. Two examples, *ka-no chi* ‘that place’ and *ka-no yumei-na NP* ‘that famous NP’ are shown below.

- (7) a. ka no chi
distant location GEN place
'that place (which is very far and hard to reach)'
- b. Taro-wa ka no yumei-na koneru daigaku-ni itta.
Taro-TOP distant location GEN famous Cornell University-DIR went
'Taro went to that famous Cornell University (that is prestigious and psychologically distant).'

Like the 1st and 2nd person pronouns, the uses of the 3rd person pronouns are also socio-culturally constrained. They are not used when referents are older than, or socially superior to, speakers, similar to the restrictions on 2nd person pronoun use. For instance, a professor could say *kare* ‘he’ or *kanojo* ‘she’ when referring to his/her student, but a

¹² ‘Woman’ in Japanese is *onna* (女). ‘She’ in Japanese was initially *ka no onna*. Around 1910, the pronunciation *onna* was replaced by *jo*, and *ka no onna* turned to *ka no jo* (Ishiguro, 2013), *jo* being another (Sino-Japanese) pronunciation of the Japanese character for *onna* (女).

student would not say *kare* ‘he’ or *kanojo* ‘she’ when referring to his/her professor. The avoidance of *kare* ‘he’ and *kanojo* ‘she’ in reference to social superiors and elders reflects a language change of the original word *ka/kare*. Where the words were formerly used as a sign of politeness toward referents, they are now used in referring to socially lower or younger people. Hinds’ (1975) questionnaire research confirmed that native Japanese speakers accepted the uses of *kare* ‘he’ and *kanojo* ‘she’ when referring to someone younger than oneself and socially inferior to the speaker. However, it is not simply the case that *kare* ‘he’ and *kanojo* ‘she’ are avoided when referents are older than or socially superior to speakers. The personal relationship between speakers and referents also affects the usage of *kare* ‘he’ and *kanojo* ‘she’. Hinds (1975) found that they are avoided when referents are personally close to speakers (e.g., family members, close friends) as well as when referents have little personal relationship to speakers and listeners (e.g., strangers, celebrities), even if the referents are younger than and socially inferior to the speaker. In sum, *kare* ‘he’ or *kanojo* ‘she’ are only used to refer to those who are younger than the speaker, socially inferior to the speaker, personally acquainted with the speaker/listener but not socially close to them. The interaction of these factors (i.e., age, social status, and personal relationship) is roughly depicted in the graphic below. The x-axis and y-axis express the status of referents relative to that of the speakers. Referents who fit in the gray rectangle can be referred to using *kare* ‘he’ and *kanojo* ‘he’.

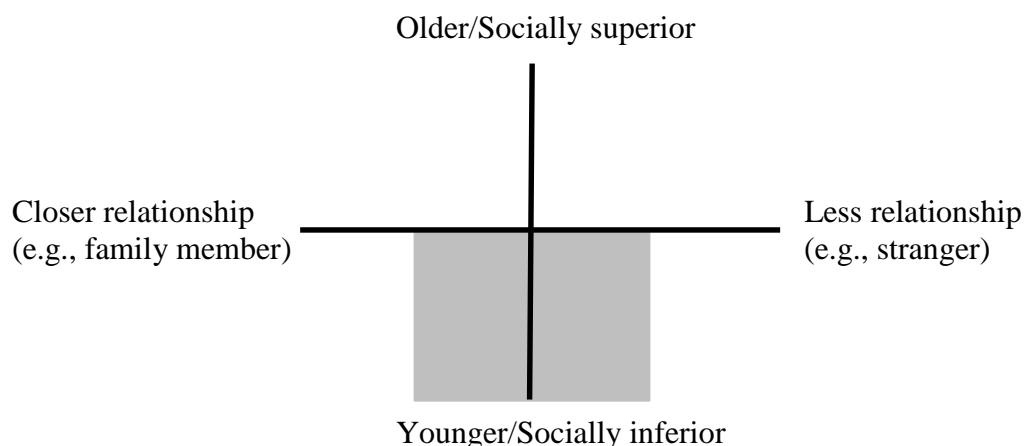


Figure 2.1. Status and familiarity of individuals who can be referred to using *kare* ‘he’ or *kanojo* ‘she’

In addition to factors listed above, Hinds (1975) also reports that Japanese speakers feel that the use of *kare* ‘he’ and *kanojo* ‘she’ is sophisticated and that they therefore tend to use *kare* ‘he’ and *kanojo* ‘she’ in referring to a foreigner. In addition, these overt pronouns are not used when the referent person is present at the time of utterance.¹³

Note that the socio-cultural constraints depicted above often do not apply in written Japanese, especially in novels (Obana, 2003b). In novels, a writer has total control over the characters, so he/she is socially neither superior nor inferior to them. Also, he/she knows the characters, but not personally. Thus, novelists often use *kare* ‘he’ and *kanojo* ‘she’ to refer to the characters (except in conversations between characters)

¹³ *Kare* ‘he’ and *kanojo* ‘she’ were also developed to form regular nouns that mean ‘boyfriend’ and ‘girlfriend’, which Hinds calls the *lover effect* (Hinds, 1975, p. 133). A famous example is a pop song from 1973, *Watasi-no kare-wa hidari-kiki* ‘my boyfriend is left-handed’. However, it is rare to call a boyfriend *kare*. Instead, *karesi* is the more common term for ‘boyfriend’.

without being constrained by socio-cultural factors.

As mentioned earlier in this section, Clancy (1980), Nagano (2014) and Miyaji (1964) revealed that 3rd person pronouns usage is much less prevalent than that of 1st and 2nd person pronouns. However, in written Japanese, as Obana (2003b), Ueno and Kehler (2010) point out that “[3rd person pronouns] are used more frequently than before in [Japanese people’s] daily life” (Obana, 2003b, p. 139), and that “they appear to be becoming incorporated into daily Japanese at an increasing rate” (Ueno & Kehler, 2010, p. 2057). An observation with the Kotonoha Shonagon corpus (<http://www.kotonoha.gr.jp/shonagon>) supports their view, as shown in the table below.¹⁴

¹⁴ In this corpus study, the search for *kare* ‘he’ was conducted by looking up this word’s kanji character, 彼. The numbers shown in the table exclude the following words that include this character: 何彼, 誰彼, 彼我, 彼岸, 彼此, 彼氏, 彼ら, 彼等, 彼女ら, 彼女等, 彼処, 彼程, 彼奴, 彼方, 彼某, 元彼, 今彼, 前彼, 彼所, and 海彼 (see Appendix A for a table showing readings and meanings of these words). In addition, the corpus also finds the increasing uses of *kare-ra* ‘they’, which is the plural form of *kare* ‘he’: 172 (1971-1979), 1472 (1980-1989), 5516 (1990-1999), and 9934 (2000-2005). Likewise, the corpus finds the increasing uses of *kanojo-ra* ‘they’, which is the plural form of *kanojo* ‘she’, as 2 (1971-1979), 48 (1980-1989), 127 (1990-1999), and 181 (2000-2005). A limitation of this corpus research is that it could not exclude *kare* and *kanojo* used as nouns for ‘boyfriend’ and ‘girlfriend’.

Table 2.5

Appearances of Kare ‘he’ and Kanojo ‘she’ in Books Published in 1971-2005

	1971-1979 (9 years)	1980-1989 (10 years)	1990-1999 (10 years)	2000-2005 (6 years)
<i>kare</i> ‘he’	538	5,113	19,697	32,046
<i>kanojo</i> ‘she’	407	2,618	10,552	18,021

Simultaneously, socio-cultural constraints on the use of the 3rd person overt pronouns appear to be loosening among younger generations. Obana (2003b) replicated Hinds’ (1975) questionnaire research with participants in different generations (15 to 75 years old). She finds that, compared to the participants older than 45 years old, those younger than 45 tend to accept *kare* ‘he’ and *kanojo* ‘she’ used in reference to family members and to both strangers and celebrities. They also tend to freely use *kare* ‘he’ and *kanojo* ‘she’ in a formal speech or writing, irrespective of the social relationship between the speaker and the referent. These findings indicate that 3rd person overt pronouns usage may be in a process of diachronic change. They might be becoming less constrained by the socio-cultural factors, and on their way to becoming more pronoun-like in the sense of Western languages.

Japanese Overt Pronoun as Epithet

The preceding sections have established that Japanese overt pronouns function as

more than simply pronouns, contrasting with Western languages. This begs the question of whether they should be categorized as overt pronouns if their use extends beyond that of overt pronouns in other languages. Based on the distributions of Japanese overt pronouns, linguists such as Hashimoto (1948), Kuroda (1965), Wetzel (1994) and Kanaya (2002) argue that there are no differences in morphosyntactic features between overt pronouns and regular nouns in Japanese, and they suggest that what is called overt pronouns should be classified as regular nouns. On the other hand, Hoji (1991), Hasegawa (1995), Kinsui (1989), Okamura (1972) and Takubo and Kimura (1992) categorize *kare* ‘he’ and *kanojo* ‘she’ as demonstratives because they take *kare* ‘he’ and *kanojo* ‘she’ to be used as indexing someone from the speaker’s viewpoint. Japanese has a rich three-position system of deictic expressions, sometimes referred to as the so-called *ko-so-a-do* paradigm. Expressions starting with *ko* indicate ‘something or someone close to the speaker’; those with *so* indicate ‘far from the speaker and close to the hearer’; those with *a* indicate ‘far from both the speaker and hearer’; and those with *do* are interrogatives. Examples are shown below.

(8)

<i>kore</i> ‘this’	/	<i>koko</i> ‘here’
<i>sore</i> ‘that (close to the hearer)’	/	<i>soko</i> ‘there (close to the hearer)’
<i>are</i> ‘that (far from speaker and hearer)’	/	<i>are</i> ‘over there (far from speaker and hearer)’
<i>dore</i> ‘which’	/	<i>doko</i> ‘where’

Hoji (1991) discusses the similar distribution of those with *a* and *kare* ‘he’ / *kanojo* ‘she’, i.e., the words with *ka*, suggesting that they are demonstratives.

Yashima (2014) presents an alternative analysis which suggests that Japanese

pronouns should be classified as epithets, a category that Jackendoff (1972) calls *specialized pronouns* (p. 110). Epithets, exemplified in English by expressions such as *the bastard*, *the idiot* or *the fool*, carry semantic connotations that are often negative. According to Dubinsky and Hamilton (1998), English epithets have nearly the same distribution as pronouns from a syntactic perspective, being subject to Condition B of the Binding Theory (Chomsky, 1981). This distribution is distinct from that of pronouns by virtue of their having the added property of being anti-logophoric. Based on the sense of the term *logophoricity* as introduced by Hagege (1974), logophoric pronouns are pronominal expressions that exclusively refer to antecedents “whose speech, thoughts, feelings, or general state of consciousness are reported” (Clements, 1975, p. 141). Logophoricity thus entails a speaker’s situating himself in the place of (i.e., empathizing with) the referent. Anti-logophoricity plays the opposite role; anti-logophoric pronouns tend not to refer to antecedents when the sentence reports the antecedents’ speech or thought (i.e., when the antecedent is the perspective-bearer). Dubinsky and Hamilton claim that the pronominal quality of epithets is supported by the fact that they are subject to Condition B of the Binding Theory. According to Condition B, pronouns may not be bound in their governing category. Accordingly, the sentence *John_i embarrassed the idiot_i* (in which *John* coreferential with *the idiot*) is ungrammatical for the same reason that *John_i embarrassed him_i* is, namely the subject *John* binds (i.e. c-commands and is co-indexed with) the object expression *the idiot/him* in its governing category (Dubinsky & Hamilton, 1998, p. 688). Further, in addition to being constrained in their distribution in the same way as regular pronouns, the distribution of epithets is also restricted by anti-logophoricity, as shown in English examples below.

- (9) a. *It was said by John_i that the idiot_i lost a thousand dollars on the slots.
 b. It was said of John_i that the idiot_i lost a thousand dollars on the slots.
- (10) a. *According to John_i, the idiot_i is married to a genius.
 b. Speaking of John_i, the idiot_i is married to a genius.

(Dubinsky & Hamilton, 1998, p. 688)

The sentences (9a) and (10a) are ruled out because the above sentences report John's (i.e. antecedent) statements, and the epithets anti-logophorically avoid referring to it. Notice that in neither (9a) nor (10a) can it be said that *John* binds the epithet, since it does not c-command it. Thus, the acceptability of the sentences in (9) and (10) can be seen to vary solely on account of the sentence taking (or not taking) John's perspective. If the sentence is logophoric relative to *John* (e.g., (10a)), then an epithet referring to *John* is unacceptable, the epithet being anti-logophoric. Yashima (2014) presents a paradigm similar to this one, claiming that Japanese 3rd person overt pronouns are also pronouns subject to Condition B of the Binding Theory, and further constrained by anti-logophoricity.

- (11) a. *Dono gakusei_i-mo kare_i-o semeta.
 every student-PART he-ACC blamed
 'Every student blamed himself'
- b. *Dono gakusei_i-mo [kare_i-ga isya-ni naru] to sinziteiru.
 every student-PART he-NOM doctor-as become that believe
 'Every student believes that he will become a doctor.'

(Yashima, 2014, p. 1430)

The sentence (11a) is ungrammatical because it violates Condition B, and (11b) is impossible because it violates anti-logophoricity constraint, i.e., *kare* 'he' is referring to

gakusei ‘student’ that is the agent of believing.¹⁵ We can observe that (11b) is not ruled out by Condition B, since *kare* is bound by an expression *dono gakusei* that is outside its governing category. The following data from Kuno (1986) also demonstrates the anti-logophoric property of Japanese overt pronouns.

- (12) a. *[Hanako_i-ga tensai da] to kanozyo_i-ga omotte iru.
 Hanako-NOM genius is that she-NOM thinking
 ‘Hanako thinks that she is a genius.’
- b. [Hanako_i-ga tensai da] to kanozyo_i-no ryoosin-ga omotte iru.
 Hanako-NOM genius is that she-GEN parents-NOM thinking
 ‘Hanako’s parents think that she is a genius.’
- (Kuno, 1986, p. 42)

The sentence (12a) is not acceptable because the sentence reports Hanako’s thought, and the overt pronoun *kanozyo* (or *kanojo*) ‘she’ is referring to *Hanako* (i.e., *Hanako* is the perspective-bearer). On the other hand, (12b) is acceptable because the sentence reports the thought of Hanako’s parents’, not Hanako’s (i.e., Hanako’s parents are the

¹⁵ The unacceptability of (11b) could be accounted for by a widely accepted view that Japanese overt pronouns cannot be interpreted as bound variables (Saito & Hoji, 1983). However, according to Yashima, this account is not enough because there are sentences in which an overt pronoun can become a bound variable, as shown in the one below.

Dono nooberysyoo zyusyoo sakka-mo kare-no hisyo-o turetekita.
 every Nobel.Prize winning author-PART he-GEN secretary-ACC brought
 ‘Every Nobel Prize winning author brought his secretary.’

(Hoji, Kinsui, Takubo & Ueyama, 2000, p. 142)

perspective-bearer). Again, neither occurrence of *kanozyo* violates Condition B, since *Hanako* does not c-command *kanozyo*.¹⁶

On the other hand, Kanno (1997) and Pimentel's (2014) questionnaire-based experiments present data in which *kare* 'he' and *kanojo* 'she' might be taken to not be anti-logophoric. They used items similar to (12a) above, such as *Yamamoto-san-ga [asita kare-ga butyooo-ni au to] itteimasu yo* 'Mr. Yamamoto is saying that he (*kare*) will meet the division manager tomorrow' (Kanno, 1997, p. 283; Pimentel, 2014, p. 174).¹⁷ For this type of sentence, Kanno and Pimentel asked native Japanese speakers whether *kare*'s antecedent was (i) *Yamamoto* (logophoric antecedent), (ii) another person (anti-logophoric antecedent) or (iii) either *Yamamoto* or another person. Their results are shown below.

¹⁶ Yashima (2014) maintains that a sentence such as (12a) is acceptable because in his view, Japanese 3rd person overt pronouns' anti-logophoricity appears only when the antecedents are quantified (e.g., *Dono gakusei-mo* 'every student' in (11a)). However, Kanno and Pimentel's data implies that the anti-logophoricity constraint extends to non-quantified antecedents.

¹⁷ A difference between (12a) and Kanno and Pimentel's sentence is that (12a) is a scrambled sentence (OSV), while Kanno and Pimentel's sentence is in the default word order in Japanese (SOV).

Table 2.6

Antecedents of kare in 'Yamamoto is saying that kare will meet the division manager tomorrow.' (Kanno, 1997; Pimentel, 2014)

Antecedent	(i) Yamamoto	(ii) Another person	(iii) Either
Kanno (1997)	7%	53%	40%
Pimentel (2014)	21%	59%	20%

Based on the results of participants' acceptances for (i) *Yamamoto* plus (iii) either *Yamamoto* or another person (7% plus 40% in Kanno; 21% plus 20% in Pimentel), Kanno and Pimentel conclude that Japanese 3rd person pronouns can refer to logophoric antecedents (e.g., *Yamamoto*), suggesting that they are not anti-logophoric. However, a closer examination of these studies actually suggests that the anti-logophoric property of *kare/kanojo* is real. One possible reason for coming to this conclusion from their studies might lie in their presentation of experimental items. In the sentences that they provided to participants, only logophoric antecedents were explicitly presented, without any competing antecedent. The absence of anti-logophoric antecedents might well have biased participants to accept an interpretation that they might not generate independently (i.e., *kare/kanojo*'s reference to logophoric antecedents). Furthermore, even with the possible bias, their results show that the participants dis-preferred the reference to (i) *Yamamoto* (a logophoric antecedent) compared with (ii) another person (anti-logophoric antecedent): 7% vs. 53% in Kanno; 21% vs. 59% in Pimentel. I interpret this strong dis-

preference for a logophoric antecedents (even in the condition wherein it is explicitly offered as a grammatically well-formed possibility) as an indication of the anti-logophoric property of Japanese 3rd person pronouns *kare* ‘he’ *kanojo* ‘she’.

As shown in the observations above, the distribution of Japanese overt pronouns share the properties with English epithets. They are pronouns, but *specialized pronouns* that do not refer to the antecedents whose speech or thoughts are reported. I use the term “Japanese overt pronouns” in this dissertation for expository purpose, but they are actually epithets and greatly different from overt pronouns in English or Spanish.

For the examination of the OPP in the current study, among the Japanese overt pronouns, I will use the 3rd person overt pronouns and will not use the 1st and 2nd person overt pronouns. The reasons are outlined below. First, to my knowledge, all the existing RNP and OPP studies focus on the 3rd person overt pronouns. Second, *kare* ‘he’ and *kanojo* ‘she’ are often used without socio-cultural connotations in written Japanese. Third, 1st and 2nd person pronouns have a numerous number of forms, whose usage is constrained by a great variety of socio-cultural factors. Managing the many overlapping conditions governing the distribution of the various 1st and 2nd person forms would likely be impossible within an experiment. On the other hand, the 3rd person overt pronouns have only a few forms, and are becoming less constrained in their use, especially among younger speakers (Obana, 2003b). Thus, the 3rd person pronouns are likely to be the best category to treat as “overt pronouns” in OPP experiments in Japanese. However, when discussing the results, we should keep in mind that the Japanese overt pronouns are not perfectly comparable to Spanish or English overt pronouns because they are anti-logophoric epithets. Experimental outcomes might therefore be subject to these effects

even if the experiments elicit OPP-like reading-time differences between overt pronouns and null pronouns.

Null Pronouns and Repeated Names in Japanese

Since overt pronouns are infrequently used in Japanese, alternative anaphors are null pronouns and repetitive full NPs, as Martin (1975) states in the following:

[...] the appropriate translation of English pronoun is either zero (omit the reference) or a repetition of the noun. In English we avoid repeating a noun once it has been mentioned, substituting an anaphoric pronoun after the first mention.

In Japanese there is no stricture against repeating the noun any number of times

[...] (p. 1075)

Kuroda (1965) and Kameyama (1985) maintain that null pronouns in Japanese correspond to English overt pronouns. Because I have shown that Japanese overt pronouns are actually epithets, null pronouns should be the only regular pronouns in Japanese. Nakayama (1990a, b) examined reference resolution with null pronouns and overt pronouns in Japanese in sentences such as the ones below; the possible antecedents and anaphors are underlined in these sentences.

(13)

a. Null pronoun

Toshokan-de	benkyooshiteiru	<u>gakusei-ga</u>	tomodachi-ni	<u>Ø</u>	shukudai-o	shiteoita
library-at	studying	student-NOM	friend-to		homework-ACC	did
to		tsugeta.				
COMP		informed				

‘The student who was studying at the library informed his friend that he did homework.’

b. Overt pronoun

Machi-o aruiteita obasan-ga keisatsu-ni kanojo-ga doroboo-o mita to
street-ACC walking woman-NOM police-to she-NOM thief-ACC saw COMP

denwashita.
telephoned

‘The middle-aged woman who was walking on the street telephoned the police that she saw the thief.’

(Nakayama, 1990b, p. 15)

Both the antecedents and anaphors are the subjects in the matrix clauses and embedded clauses, respectively. Nakayama employed a probe recognition task in this experiment. After participants read the experimental sentences, they saw probe words that were the antecedent NPs and judged if the words were present in the sentences that they just read. The results show that the participants’ reaction times for the null pronouns’ antecedents were significantly faster than those for the overt pronouns’ antecedents. This outcome indicates that the participants’ awareness of the referential relationship between null pronouns and antecedents was faster than their awareness of the relationship between overt pronouns and antecedents. This result suggests an effect similar to the OPP is operative in Japanese. That is, sentences with overt pronouns would be read slower than those with null pronouns. However, in order to confirm this implication, it is necessary to further test the reference to non-salient antecedents (e.g., object antecedents) and see whether or not there is a significant difference in reading-time between null-pronoun and overt-pronoun sentences. In addition, as mentioned in the previous section, it is important to note that Japanese overt pronouns are different from Spanish overt pronouns. The slower reading times of Japanese overt pronouns might be due to their epithet-like quality. It is possible that the participants’ establishment of reference in (13b) was slower because

kanojo ‘she’ would be unlikely to refer to *obasan* ‘middle-aged woman’ in this sentence when *obasan* ‘woman’ is the agent of reporting (i.e., *denwashita* ‘called’). Moreover, since the use of Japanese overt pronouns could sound unnatural due to socio-cultural constraints such as antecedents’ ages or social statuses, it could be more appropriate not to use kinship terms or title, e.g., *obasan* ‘middle-aged woman’.

In investigating the distribution of null pronouns and full NPs, a corpus study by Hashimoto, Inui, Shirai, Tokunaga and Tanaka’s (2001) found that null pronouns refer to antecedents that are closely positioned to the null pronouns, while a repetitive NP is often used in a new paragraph, referring to the antecedents in a previous paragraph. Also, Tsuchiya, Yoshimura and Nakayama (2014) report that Japanese speakers use an anaphoric full NP when another NP interrupts the referent and the full NP. Similarly, Yoshida’s study (2005a, 2005b, 2011) utilizes Walker’s (1998) Cache Model, which introduces the notions of *cache*, *push*, and *return pop*. Caches are the current centers of attention, which are active in short-term memory and are easily recoverable entities; pushes are newly activated centers of attention; and return pops are reactivated centers of attention. According to Yoshida, a cache is initially introduced as a full NP and maintained by null pronouns. This entity is considered to be the current topic and currently activated within a scene, an episode, a paragraph, etc. The maintenance of cache (i.e., the current topic) can be interrupted by another full NP, push, which expresses a subtopic (i.e., a new center of attention). After a push cut the continuity/chain of the cache, full-NP anaphors are used as a return pop in order to resume the cache. This is illustrated in the example below.

(14)

1. 'Taro_i studied at the library.' ← full NP_i = introduction of a topic (cache)
2. '∅_i studied physics.' ← null pronoun_i = reference to the topic
3. 'Hanako_y came into the library.' ← full NP_y = introduction of a subtopic (push)
4. 'Taro_i talked to her.' ← full NP_i = resumption of the topic (return pop)

Nakahama's (2011) observations regarding Japanese speakers' narratives support the above analyses. Nakahama found that a full NP was used to introduce a new entity, which was referred to by null pronouns. A reference by a full NP occurred when the speakers re-introduced the topic after another subtopic intervened. The studies above agree that null pronouns refer to a currently activated antecedent while repeated full NPs refer to an inactivated information (i.e., new entity) or less activated information (e.g., non-salient antecedents that appeared in the previous paragraph). These uses of Japanese full NPs seem similar to English full NPs, as Marslen-Wilson, Levy and Tyler (1982) found that English speakers tend to use full NPs when re-establishing a non-salient referent into a central role of the discourse.

Chapter 3. Adnominal Postpositions, *Wa* and *Ga*

Wa and *Ga* for Subject

So far, this paper has discussed the forms of anaphoric expressions (i.e., null pronouns, overt pronouns and repeated full NPs). However, Japanese utilizes adnominal postpositions with NPs to signal grammatical and semantic roles as well as to mark discourse functions. It should be noted that there is not a one-to-one relationship between form and meaning, or between form and function. A single grammatical function or semantic role can be marked by more than one postposition and particular postpositions can be used to signal more than one function or grammatical role. For instance, the subject in Japanese is typically appears with the topic-postposition *wa* and the nominative-postposition *ga*. Recall that the previous studies regarding the RNP and OPP detected slower reading times when the anaphors are grammatical subjects. Unlike English or Spanish, the difference of reading times of Japanese anaphors should be considered not only with reference to different forms (i.e., null pronouns, overt pronouns and repeated names) but should also take into account which postpositions are attached to them. The subject in Japanese is typically (i) a null topic-NP, (ii) an overt topic-NP marked by *wa*, or (iii) an overt non-topic subject-NP marked by *ga*. In addition, the overt NPs can be either repeated names or overt pronouns (e.g., ‘John-*wa*’ or ‘He-*wa*’, and ‘John-*ga*’ or ‘He-*ga*’). Thus, the subject anaphors can have five common variations: (i) null pronoun, (ii) repeated NP-*wa*, (iii) repeated NP-*ga*, (iv) overt pronoun-*wa*, and (v) overt pronoun-*ga*. The current study utilized these five types of anaphors. The current

section specifically reviews the functions of the topic-postposition *wa* and the nominative-postposition *ga* and relates them to anaphoric expressions.

Topic *Wa* and Non-topic *Ga*

Many studies have discussed the informational differences between sentences containing NP-*wa* and NP-*ga*. The most well known argument on this issue is that the NP-*wa* is a topic that refers to information that is already acknowledged (i.e., activated in memory), while the NP-*ga* is a non-topic that refers to unacknowledged information (Kasuga, 1918; Matsushita, 1928; Kuno, 1973a, b; Hinds, 1983; Iwasaki, 1987; Shibatani, 1990). While Kuno (1973a, b) states that an NP-*ga* cannot refer to old information (unless it focuses on an antecedent), Noda (1996) argues that, although an NP-*ga* could refer to an antecedent, the sentence with an NP-*ga* has to be discontinuous, describe unpredictable events, or change the scene in the discourse.

The following two question-answer exchanges demonstrate the use of *wa* for an acknowledged topic-NP and *ga* for an unacknowledged non-topic NP.¹⁸

- [illegible]

¹⁸ The parentheses in the examples mean that these NPs can be null because they are the topics, which can be expressed in the form of null pronouns, as mentioned earlier. Also, in (15b), because Hanako is the topic (i.e., old information), it could be marked with *wa*, as *Hanako-wa Taro-ga tataita* ‘Hanako, Taro slapped her’.

In discourse (15a), Taro must be appended with *wa* (or null) because it is acknowledged information (i.e., topic) from the preceding question sentence. On the other hand, in (15b), Taro must be appended with *ga* because it is not acknowledged information within the surrounding context.

Although Japanese explicitly expresses the topic-NP in the null form or with the topic-postposition *wa*, the discourse role of topic is universal - even though it might be realized through differing morphosyntactic means; this includes English and Spanish, which do not utilize adnominal postpositions. For example, in the English discourse, “I saw John yesterday. He was walking down the street.”, *He* is the topic, whose identity is already acknowledged before the second sentence is even uttered (i.e., *He was walking...* is equivalent to *John was walking...*) because its referent appears in the preceding sentence (i.e., *I saw John yesterday*). The topic is “what the sentence is about” (Gundel, 1988, p.17), such that the phrase “walking down the street” is describing about *John/He*. Weil (1869/1978) analyzes sentences possessing both a topic and its description in the following manner: a speaker starts his or her utterance at the point of departure, and he or she then adds comments to the rest of the utterance. The point of departure contains the information which the speaker assumes the hearer already knows (i.e. topic, acknowledged information); the added comments are what the speaker assumes the hearer does not know and, likewise, what the speaker wishes to convey to the hearer. Prince (1978) states that a speaker cannot mark an entity as the topic unless the hearer is already familiar with the entity that the speaker is going to talk about. In response, the speaker uses the entity marked as the topic to lead the hearer back to the antecedent in the preceding context, in the ongoing situation, or in his or her long-term memory (Haviland

& Clark, 1974). These analyses of the topic as already acknowledged information suggest that the topic should be often used as an anaphor that refers back to previously established information (i.e., antecedent).

The topic is analyzed in the framework of the information structure, as opposed to the syntactic structure that consists of the subject, object, etc. Weil (1869/1978) and Mathesius (1899/1961) argue that sentences should be analyzed in two different linguistic frames: grammatical structure (i.e., syntactic structure), relating to objective description and consisting of the subject and predicate, and functional structure (i.e., information structure), which is based on speaker's subjective observations and consists of the topic and comments about it. Halliday (1967) first coined the term *information structure* (p. 199) for the functional structure mentioned by Weil and Mathesius. The differentiation of the information structure (i.e. topic-comment) and syntactic structure (i.e. subject-predicate) is widely accepted because "there is often no one-to-one correspondence between syntactic form and discourse function" (Lambrecht, 1994, p. 32), and "it is impossible for grammars to directly mirror any one functional motivation, since such motivations tend to conflict with each other" (Newmeyer, 1999, p. 478). In short, an NP syntactically categorized as type-X does not necessarily map to a particular discourse function Y, nor vice versa. For instance, an NP can be simultaneously the subject (as the syntactic category) and the topic (as the information-structural category), but an NP can be also the object (as the syntactic category) and the topic (as the information-structural category). English examples of these various topics are shown below.

(16) a. Subject = Topic

Q: Who did John slap? A: He slapped Mary

b. Object = Topic

Q: Who slapped John? A: Mary slapped him.

c. Locative phrase = Topic

Q: What happened at John's house? A: Chris slapped Mary there.

The topic *He* (i.e., *John*) is the syntactic subject in (16a). The topic *him* (i.e., *John*) is the object in (16b). Finally, the topic *there* (i.e., *John's house*) is the locative prepositional phrase in (16c). As shown in the examples above, the topic in English is commonly expressed with overt-pronoun anaphors, which is related to the RNP found by Gordon, Grosz and Gilliom (1993). Thus, overt pronouns are preferred to non-pronouns (e.g., repeated names) when referring to already acknowledged information, except the case where there are multiple possible antecedents. However, Gordon et al. found that the preference for overt pronouns is strong especially when the pronoun is the subject; a non-pronoun subject anaphor causes slowed reading, but a non-pronoun non-subject anaphor does not elicit this penalty.

In Japanese, the topic is expressed in the form of an overt NP with *wa* or a null pronoun with no affixes. If the topic necessarily refers to acknowledged information or if an anaphor is essentially the topic, null pronouns and the topic-NP-*wa* are more prone to be used as an anaphor than the non-topic-NP-*ga*. According to Gordon and Hendrick's (1998) DPT, repeated names are first interpreted non-anaphorically, which slows down reading and elicits the RNP. If Japanese topic-NP-*wa* is immediately interpreted as an anaphor because of the topic-postposition *wa*, then it should be read slower for salient

antecedents than for non-salient antecedents, similar to null pronouns (even if the topic-NP-*wa* is a repeated name), resulting in no RNP. On the other hand, Almor's (1999) ILH suggests that the rich semantic information in a repeated name delays the reading speed compared to overt pronouns (in English) and null pronouns (in Spanish). If this is true, even with *wa*, reading times of repeated-name anaphors and null pronouns should show interaction with salient/non-salient antecedents (i.e., RNP). In addition, in terms of the comparison between the overt topic-*wa* and null-pronoun topic, Clancy and Downing's (1987) corpus study with Japanese speakers' spoken narratives found that null pronouns most frequently referred to subject antecedents, while overt topic-NPs with *wa* rarely referred to the subject antecedents. Hinds' (1984) interview study also found that after a new entity was introduced with a full-NP-*ga*, it was referred to by null pronouns unless there were ambiguous antecedents. Nakahama's (2003) observation of Japanese speakers' narrations also found that the most frequent antecedent-anaphor combination was a new NP with *ga* followed by a null pronoun. Almor's ILH and Clancy and Downing, Hinds, and Nakahama's studies imply that the anaphoric topic-NP-*wa* should not be used to refer to salient (subject) antecedents and would trigger the RNP. In reference to salient antecedents, such as the subject, sentences with a repeated-name-*wa* would be read slower than sentences with a null pronoun, but this slowed reading would be neutralized to some extent if antecedents are non-salient objects.

Shoji, Dubinsky and Almor (2015, 2016) recently examined the RNP and OPP with both the topic-*wa* anaphor and non-topic-*ga* anaphor as well as null pronouns in Japanese. Their results support the above prediction from Almor (1999)'s ILH. The results showed that both repeated-name topic-*wa* anaphors and repeated-name non-topic-

ga anaphors exhibited the RNP and that both overt-pronoun topic-*wa* and overt-pronoun non-topic-*ga* showed the OPP. In other words, regardless of the topic-hood of the anaphors, both repeated names and overt pronouns were read significantly slower than null pronouns when referring to subject antecedents. These reading-time differences were significantly decreased when those anaphors were referring to object antecedents. These results support Almor's ILH, but not Gordon and Hendrick's DPT.

Exhaustive-listing *Ga* and Contrastive-topic *Wa*

Although the non-topic subject-NP-*ga* does not refer to acknowledged information in general, Kuno (1973a, b) discusses that the NP-*ga* can be anaphoric when *ga* exhaustively lists the NP. The function of exhaustive-listing *ga* is to express, "X, and only X among all others under discussion" (Mikami, 1963; Kuno, 1973a, b). An example is shown below.

- (17) Taro to Hanako to Ziro-ga toshokan-ni itta.
 Taro and Hanako and Ziro-NOM library-DIR went
 'Taro, Hanako and Ziro went to a library.'
- Taro-ga hon-o yonda.
 Taro-NOM book-ACC read
 'Taro (not Hanako or Ziro) read a book.'

This usage of *ga* focuses on *Taro*, indicating that it was *Taro* who had read the book, not *Hanako* or *Ziro*. Thus, a precise English translation of a sentence with exhaustive-listing *ga* should be the it-cleft sentence in the form of 'It is X-*ga* that...' (e.g., 'It was Taro who read a book').

Similarly, in some contexts, the NP-*wa* can focus on the NP by selecting it from a group of possible antecedents. This usage is called the contrastive *wa* (Mikami, 1963; Kuno, 1973a, b). Contrastive *wa* expresses that the NP-*wa* does an action or has attained

a status, while the other possible antecedents' actions or statuses could be unknown, the same as, or different from NP-*wa* (Hara, 2006; Heycock, 2008; Deguchi, 2009; Wang & Schumacher, 2013). An example discourse is shown below.

(18)

1. Taro to Hanako to Ziro-ga toshokan-ni itta.
 Taro and Hanako and Ziro-NOM library-DIR went
 'Taro, Hanako and Ziro went to a library.'

- 2a. Taro-wa hon-o yonda.
 Taro-CONT hon-ACC read
 'Taro read a book.'
 (The speaker knows that Taro read a book, but he/she don't mention the others.)

- 2b. Taro-wa hon-o yonda ga Hanako to Ziro-mo yonda.
 Taro-CONT hon-ACC read but Hanako and Ziro-also read
 'Taro read a book, but Hanako and Ziro also did.'

- 2c. Taro-wa hon-o yonda ga Hanako to Ziro-wa yomanakatta.
 Taro-CONT hon-ACC read but Hanako and Ziro-CONT didn't read
 'Taro read a book, but Hanako and Ziro didn't.'

- 2d. Taro-wa hon-o yonda ga Hanako to Ziro-wa benkyo-sita.
 Taro-CONT hon-ACC read but Hanako and Ziro-CONT studied
 'Taro read a book, but Hanako and Ziro studied.'

All of the sentences marked 2 (a-d) show that the speaker perceives *Taro*, *Hanako* and *Ziro*. In other words, *Taro-wa* in the second sentence retrieves all three as a set of antecedents from the first sentence (i.e., 'Speaking of Taro, Hanako and Ziro, ...'). However, the speaker focuses on *Taro* among the three possible antecedents and implicitly (as in 18-2a) or explicitly (as in 18-2b, 2c, 2d) contrasts *Taro* with the other members. In the implicit case, the speaker is overtly referring to *Taro* and covertly referring to *Hanako* and *Ziro*. The speaker only mentions that at least *Taro* read a book, and he or she (intentionally or unintentionally) does not mention *Hanako* and *Ziro*

because the speaker does not want or need to talk about them or does not know what they did. Thus, an NP with the contrastive *wa* is the topic in the sense that it refers to (a set of) antecedents, and it is simultaneously the focus in the sense that it focuses on one of the possible antecedents (i.e., ‘Speaking of X, Y and Z, X-*wa*...’).¹⁹

Exhaustive-listing *ga* and contrastive *wa* are similar to English stressed pronouns and Spanish overt pronouns in that they all focus on an NP. For the purposes of this dissertation study, a potential problem is that an NP-*ga* could be interpreted either as a simple non-topic subject (i.e., ‘X...’) or as an exhaustive listing (i.e., ‘Only X, not Y or Z,

¹⁹ Although both exhaustive-listing *ga* and contrastive *wa* function to focus on an entity, they are different. An NP with exhaustive-listing *ga* indicates that this NP is the only one which/who does a certain action or has attained a certain status and that the other members do not do the action or has not attained the status. On the other hand, an NP with contrastive *wa* only indicates that the NP-*wa* does the action or has attained the status, and it remains unclear if the other members under discussion do or do not do the same action or has attained the same status. Thus, unlike a sentence with exhaustive-listing *ga* that allows the English translation in the form of it-cleft, a sentence with contrastive *wa* does not allow it. The following sentence might well depict the difference of the exhaustive-listing *ga* from the contrastive-topic *wa*.

Hanako-de-wa-naku	Taro-ga/*wa	hon-o	yonda.
Hanako-not	Taro-NOM/CONT	book-ACC	read
‘Taro, not Hanako, read a book.’			

The phrase *Hanako-de-wa-naku* ‘not Hanako’ in the sentence above overtly mentions the other member, indicating that *Taro* read but *Hanako* did not read. In this case, *Taro-ga* should be used and is not replaceable with *Taro-wa*.

...'), and that an NP-*wa* could be interpreted either as a simple topic (i.e., 'Speaking of X, X...') or as a contrastive topic (i.e., 'Speaking of X, Y and Z, X...').²⁰ That is, readings of NP-*ga* and NP-*wa* will not be slowed if readers focus the NP in their reading, while they will be slowed and will then elicit the RNP (and OPP) if readers read them as unfocused NPs. In short, the reading speeds of sentences including an NP-*ga* and those including an NP-*wa* can be inconsistent depending on whether readers interpret the former as a non-topic subject or as an exhaustive list, and the latter as a topic or as a contrastive topic. This possibility might not be problematic in the standard RNP/OPP experimental design, considering that the participants of Gordon, Grosz and Gilliom (1993) and Gelormini-Lezama and Almor (2011) showed the RNP and OPP, which indicates that they did not interpret the English and Spanish repeated names and Spanish overt pronouns as focused entities; if they did, all these anaphors should not have been read slower than reduced forms of anaphors (i.e., English overt pronouns, Spanish null pronouns), so discourses such as *Ada conoció a Nicolás. Ada/Ella lo detesta.* 'Ada met Nicolás. Ada/She hates him.' and *Ada conoció a Nicolás. Ø Lo detesta.* 'Ada met Nicolás. Ø hates him.' would not show an interaction between anaphors and antecedents in the reading times of the second sentences. However, in order to make sure that the current study in Japanese would avoid participants' exhaustive-listing interpretations of *ga* and contrastive interpretations of *wa*, the experimental items were created in such a way that the items would not allow these interpretations. For instance, a discourse, 'Taro hired Hanako as a part-timer. Taro-*ga/wa* worked with Hanako everyday since then.'

²⁰ Both exhaust-listing *ga* and contrastive *wa* are pronounced with phonological stress, but in the current study, the stimuli sentences are given in writing.

wouldn't allow exhaustive-listing/contrastive interpretations. An exhaustive-listing/contrastive interpretation such as, 'Taro worked with Hanako every day since then (but Hanako didn't / but it is unknown what Hanako did)' would not make sense because if Taro worked with Hanako then Hanako must have also worked with Taro. This way the current study excludes exhaustive-listing and contrastive interpretations and concentrates on the non-topic NP-*ga* and topic NP-*wa*, so that this study sticks to the line of the earlier RNP and OPP studies such as Gordon, Grosz and Gilliom (1993) and Gelormini-Lezama and Almor (2011).

Chapter 4. Antecedents

Antecedent Salience, the Repeated Name Penalty, and the Overt Pronoun Penalty

As shown in the series of studies related to the RNP and OPP (Gordon, Grosz & Gilliom, 1993; Gelormini-Lezama & Almor, 2011; Almor, de Carvalho Maia, Cunha Lima, Gelormini-Lezama & Vernice, 2013), it has been shown that the salience of the antecedents affects the reading speed of the anaphors. The RNP and OPP were detected only when the antecedents were salient. Salience of the antecedents refers to the activation of the information in readers' memory. It is associated with the grammatical position of the subject and may be related to where the antecedent is mentioned in the sentence (i.e. word order). For highly activated (i.e., salient) antecedents, the reading time of subsequent sentences with English pronouns were significantly faster than those with English repeated-name anaphors. In contrast, for relatively less salient antecedents (e.g. object), reading times of sentences with pronouns and repeated names were not significantly different (Gordon et al, 1993). Relatedly, reading times of sentences with Spanish null pronouns were faster than those with overt pronouns and repeated names when they refer to grammatical subjects, which are salient antecedents. However, reading times of sentences with null pronouns were slower than those with overt pronouns and repeated names when they refer to non-salient antecedents (e.g., objects) because their null pronouns were not as informative (Gelormini-Lezama & Almor, 2011).

Antecedent Accessibility

The findings from the previous studies indicate that the more salient an antecedent is, the more accessible the antecedent is; thus readers can build reference with less informative anaphors. Here, accessibility refers to the level of ease with which hearers retrieve the antecedent's information in their memory. Ariel's (1990) Accessibility Theory maintains that speakers choose anaphor form (e.g. repeated name, pronoun) depending on the accessibility of antecedents. According to Ariel, pronouns are used when the antecedents are salient and therefore highly accessible, while repeated names are used when the antecedents are not salient and minimally accessible. Gundel, Hedberg and Zacharski's (1993) Givenness Hierarchy also explains that the speaker's choice of anaphor form signals the accessibility of antecedents, which enables addressees to restrict the set of possible antecedents. Grosz, Joshi and Weinstein's (1995) Centering Theory ranks the English antecedents' accessibility based on their grammatical categories, as shown below.

(19) Ranking of accessibility (i.e., the likelihood of an NP being referred to by an anaphor)

SUBJECT > OBJECT(S) > OTHER²¹ (p. 214)

In this ranking, the subject-NP most likely becomes an antecedent, which means that the subject-NP is the center of speakers' attention. In other words, the speakers are likely to continue to talk about the NP in their utterances after the NP is previously presented as the subject.

²¹ Grosz et al. (1995) call possible antecedents the forward-looking centers.

Crawley's (1986) experiment illustrated that subject antecedents are more accessible for subject anaphors than object antecedents. She provided English sentences with overt pronouns that could ambiguously refer to subjects or objects in preceding sentences. The task of native English-speaking participants was to decide whether overt pronouns referred to the subject or object. The results showed that subjects were referred to more often than objects. Colonna, Schimke and Hemforth's (2012) experiment in German with native German speakers also shows that the subject-hood of antecedents enhances the accessibility of the antecedents. Similar to Crawley's study, the participants had to choose between one of two possible antecedents, subject or object, to be referred to by overt pronouns (e.g., *Peter has slapped John when he was young*; Colonna et al., p. 4). Participants tended to interpret the pronouns referring to the subjects more frequently than the objects. Collectively, these studies indicate that the subject-hood of antecedents enhances the accessibility. However, Colonna et al. also conducted their experiment with native French speakers and found contradictory results. Native French speakers tended to refer to the object more frequently than the subject. This opposite result implies that the preference to choose subject antecedents might not be universal among languages.

Antecedent Accessibility and Topic

In addition to the antecedents' grammatical categories, Ariel (1990) suggests that the accessibility of an antecedent is affected by whether the antecedent is a topic or non-topic. Gundel, Hedberg and Zacharski (1993) also maintain that structural markings such as the topicalization of a referent foregrounds the referent and makes it more accessible for anaphors. Givón (1983) argues that the subject is the most likely to be the topic as shown in his scale below.

(20) The likelihood of NP arguments being the sentential topic

[most likely] Subject ---- Direct object ---- Indirect object [least likely] (p. 57)

Bellitii, Bennati and Sorace (2006) also state, “[a] lexical subject in preverbal position is normally interpreted as given, topic-like information” (p. 660). The arguments by Ariel, Gundel et al., Givón, and Bellitti et al. above suggest that the subject is ranked as the most possible antecedent (in Grosz, Joshi and Weinstein’s (1995) ranking) because the subject is most likely the topic by default.

In addition to manipulating whether an antecedent was a subject or object, Crawley (1986) also manipulated the topic-hood of the antecedents. In her experiment, she included four possible antecedents: topic in the subject position, non-topic in the subject position, topic in the object position, and non-topic in the object position, all of which were ambiguously referred to by overt pronouns. In comparing the topic antecedents (i.e., topic-subject and topic-object) and non-topic antecedents (i.e., non-topic-subject and non-topic-object), the participants tended to refer to the topics more frequently than the non-topics. This result illustrates that the topic-hood of antecedents enhances the accessibility. Colonna, Schimke and Hemforth’s (2012) experiment in German speakers also elicited the similar result with Crawley’s experiment. The antecedents in their experimental items were constructed to be either topic-subject or non-topic-subject (e.g., *As for Peter, he has slapped John when he was young* vs. *Peter has slapped John when he was young*; Colonna et al., pp. 4-5). Native German speakers referred to the topic-subjects more frequently than the non-topic subjects. However, the results from Colonna, Schimke and Hemforth’s experiment with French speakers were inconsistent with the results from the German speakers. Native French speakers did not

exhibit the same preference as German speakers (Colonna et al., 2012) or English speakers (Crawley, 1986).

Recall that Gordon, Grosz and Gilliom (1993) and Gelormini-Lezama and Almor (2011) found the slower reading of repeated names and overt pronouns with salient antecedents but not with non-salient antecedents. If topic-hood enhances antecedents' accessibility and salience, topic antecedents require an anaphor that is immediately interpreted as an anaphor (according to the DPT) or an anaphor that carries less semantic information (according to the ILH). On the other hand, non-topic antecedents accept a reference by an anaphor that is not immediately interpreted as an anaphor or an anaphor that carries rich semantic information. Accordingly, the RNP and OPP would be elicited with topic antecedents and with non-topic antecedents, although the antecedents are in the same grammatical categories. For example, in an English discourse such as "As for Peter_i, he_i (topic antecedent) has slapped Mary. He_i hated her.", the second sentence with a pronoun anaphor (*He*) may be read faster than the equivalent sentence in "Peter_i (non-topic antecedent) has slapped Mary. He_i hated her." Also, in a Spanish discourse "As for Peter_i, he_i has slapped Mary. Ø_i hated her." that includes the topic-subject antecedent, the sentence that includes a null-pronoun anaphor (Ø) may be read faster than that in "Peter_i has slapped Mary. Ø_i hated her." that includes the non-topic-subject antecedent.

Topic and Subject and Antecedent Accessibility in Japanese

A Japanese subject-NP and topic-NP might be highly salient and accessible in discourse, similar to Crawley's (1986) results in English and Colonna et al. (2012)'s in German. A number of studies of Japanese discourse tested whether the subject-hood and topic-hood of antecedents enhanced the antecedents' accessibility. Clancy and Downing's

(1987) corpus study with native Japanese speakers' narratives measured how likely topic antecedents and non-topic antecedents are referred to in immediately following sentences. If the topic-hood enhanced the salience and accessibility of antecedents, the topic antecedents would be more frequently referred to than the non-topic antecedents. However, they found that there was no clear difference in accessibility between the two types of antecedents. This result stands in contrast with results from English (Crawley, 1986) and German (Colonna, Schimke & Hemforth, 2012). In addition, Clancy and Downing (1987) also found that non-topic-*ga* antecedents were referred to by null pronouns more frequently than the cases that the topic-*wa* antecedents were referred to by null pronouns. Thus, there was no indication that the topic-hood enhances an antecedent's accessibility.

In addition, Ueno and Kehler (2010) conducted an off-line sentence-completion experiment using overt and null pronouns. They manipulated the possible antecedents in their experiment. The first sentence could include the topic-subject-*wa* antecedent, the non-topic-subject-*ga* antecedent, or indirect-object-*ni* antecedent. These were followed by the second sentence that included anaphors: either overt-pronoun-*wa* or null pronoun. Example discourses in their study are shown below.

(21) a. Possible antecedents: Subject-topic vs. Indirect Object

<u>Taro</u> -wa	<u>Jiro</u> -ni	hon-o	watasita.	Kare-wa/Ø ____.
Taro-TOP	Jiro-DAT	book-ACC	handed	he-TOP/Ø
'As for Taro, he passed a book to Jiro. He/Ø ____.'				

b. Possible antecedents: Subject vs. Indirect Object

<u>Taro</u> -ga	<u>Jiro</u> -ni	hon-o	watasita.	Kare-wa/Ø ____.
Taro-NOM	Jiro-DAT	book-ACC	handed	he-TOP/Ø
'Taro passed a book to Jiro. He/Ø ____.'				

In the examples above, overt and null pronouns could refer to *Taro-wa* (i.e., topic-subject), *Taro-ga* (i.e., non-topic-subject) or *Jiro-ni* (i.e., indirect object). If the subject antecedent is more accessible than the non-subjects, anaphors' reference to the non-topic-subject (i.e., *Taro-ga*) and topic-subject (i.e., *Taro-wa*) should be more frequent than anaphors' reference to the object (i.e., *Jiro-ni*). Also, if the topic antecedent is more accessible than the non-topic antecedent, anaphors' reference to the topic-subject antecedent (i.e., *Taro-wa*) should be more frequent than anaphors' references to the non-topic subject (i.e., *Taro-ga*) and object (i.e., *Jiro-ni*).

In the Japanese participants' responses, there was no difference between the references for topic-subject antecedents and for non-topic-subject antecedents, suggesting that topic-hood did not enhance the salience/accessibility of antecedents in Japanese. Meanwhile, they found that both null and overt pronouns in the participants' completed sentences more often referred to the subject and topic-subject antecedent compared to the object antecedent. This illustrates that the subject-hood enhances the antecedents' accessibility. Also, preferences for subject (both topic and non-topic) antecedents relative to object antecedents were significantly stronger with null pronouns than with overt pronouns. Ueno and Kehler (2010) suggest that what is crucial for references by null and overt pronouns in Japanese are the antecedents' subject-hood, not the topic-hood. However, note that Ueno and Kehler used only topic anaphors (i.e., null pronouns and overt-pronoun topic-*wa*) and did not use the non-topic anaphor (e.g., overt pronoun non-topic-*ga*).

Ueno and Kehler's (2010) experiment positioned the topic-*wa* at the beginning of discourses. Therefore, these topic phrases in the first sentences did not have antecedents.

Earlier in this paper, I discussed that the topic has to refer to already-acknowledged information (e.g., antecedent), which appears to contradict to Ueno and Kehler's discourse setting. However, in Japanese, especially in writing, it is common that a discourse start with a topic phrase. Gundel (1988) states that a topic-initial sentence such as the one in Ueno and Kehler's study (e.g., 'Taro-*wa* passed a book to Jiro') can be an answer to the implicit question, "What about Taro?", and that "[a topic-initial sentence] alone can be used to begin a discourse" (p. 18). A sentence-initial topic is often the center of attention in a discourse, e.g., the protagonist of a story is marked by *wa* when he/she appears at the beginning of the entire story. In this manner, the speaker or author implicitly conveys a message to the reader or listener that the discourse/story will be focus on the topic-*wa* entity. For example, a sentence such as 'Taro-*wa* passed a book to Jiro' would sound like "Our familiar Taro passed a book to Jiro." Given a topic-initial sentence, listeners or readers also regard the sentence-initial topic as a central figure in the following discourse. Novelists often start a story with a topic-NP-*wa* as a strategy so as to make readers take the referent of this NP to be an already familiar character (Noda, 1996; Maynard, 2004). The story that follows can then proceed to answer the question, "what happened to him/her?" On the other hand, a topic-less sentence such as 'Taro-*ga* passed a book to Jiro' objectively describes about *Taro* (and *Jiro*). This sentence does not present a discourse topic. According to Maynard (2004), the phrase *Taro-ga* only temporarily spotlights and foregrounds *Taro*, which is the subtopic or "push" in Walker's (1998) Cache Model term.

Like Ueno and Kehler (2010), Okuma (2011) also tested various antecedents' accessibility in Japanese. Okuma provided sentences that included either null- or overt-

pronoun subject anaphors that could refer to one of the following competing antecedents: topic-subjects, non-topic-subjects, topic-indirect-objects, or non-topic-indirect-objects. In addition, she manipulated the word order of these possible antecedents. Example items are shown below.

(22)

a. Antecedents: Topic-subject vs. Non-topic-object (topic-subject is first mentioned)

Okaasan-wa musume-ni Ø/kanojo-ga kotoo-o kiru tokini kisu-o sita.
 mother-TOP daughter-DAT Ø/she-NOM coat-ACC put.on when kiss-ACC did
 ‘As for the mother, she kissed the daughter when Ø/she put on the coat.’

b. Antecedents: Topic-object vs. Non-topic-subject (topic-object is first mentioned)

Musume-ni-wa okaasan-ga Ø/kanojo-ga kotoo-o kiru tokini kisu-o sita.
 daughter-DAT-TOP mother-NOM Ø/she-NOM coat-ACC put.on when kiss-ACC did
 ‘As for the daughter, the mother kissed her when Ø/she put on the coat.’

c. Antecedents: Non-topic-subject vs. Non-topic-object (subject is first mentioned)

Okaasan-ga musume-ni Ø/kanojo-ga kotoo-o kiru tokini kisu-o sita.
 mother-TOP daughter-DAT Ø/she-NOM coat-ACC put.on when kiss-ACC did
 ‘The mother kissed the daughter when Ø/she put on the coat.’

d. Antecedents: Non-topic-object vs. Non-topic-subject (object is first mentioned)

Musume-ni okaasan-ga Ø/kanojo-ga kotoo-o kiru tokini kisu-o sita.
 daughter-DAT mother-NOM Ø/she-NOM coat-ACC put.on when kiss-ACC did
 ‘The mother kissed the daughter when Ø/she put on the coat.’

(Okuma, 2011, p. 94)

The participants’ task was to judge which antecedent the anaphors referred to. The results showed that the participants did not show any preference for the topic status or the word order of the antecedents, consistent with Ueno and Kehler’s experiment. On the other hand, Okuma also found that participants interpreted the topic-subjects and non-topic-

subjects (i.e., *okaasan* ‘mother’) as the antecedents of the null pronouns rather than the topic-objects and non-topic-objects; however, the participants did not show a preference for any type of antecedents when the anaphors were overt pronouns.

Both studies by Ueno and Kehler (2010) and Okuma (2011) suggest the subject-hood is crucial to the antecedents’ accessibilities in Japanese, regardless of the topic-hood. None of the above studies, including Clancy and Downing’s (1987) study, found that antecedents’ topic-hood enhanced the antecedents’ accessibility. This data implies that topic antecedents and non-topic antecedents in the same grammatical categories (e.g., topic subject and non-topic subject) might be equally salient and accessible. If this were the case, then anaphors would refer to the topic and non-topic antecedents indifferently in this dissertation study and thus the RNP and OPP would not appear with the two types of antecedent. Considering Clancy and Downing’s participants used null pronouns more frequently when referring to the non-topic-subject-*ga* antecedents than when referring to the topic-subject-*wa* antecedents, it is possible that the RNP and OPP may indicate that the non-topic antecedents are more salient/accessible than the topic antecedents.

Although Ueno and Kehler’s (2010) and Okuma’s (2011) experiments above showed that the subject-hood of antecedents is more important than the topic-hood, they also found that the degree of the accessibility of subject/object antecedents differs between null pronouns and overt pronouns. Ueno and Kehler’s participants showed a weaker preference for referring to subject/topic-subject antecedents when anaphors were overt-pronoun-*wa* than when they were null pronouns, and Okuma’s participants did not

even show any antecedent preference when the overt-pronoun-*ga* were used.²² Nagano's (2014) experiment shows even more radical results. She tested the accessibility of subject and non-subject antecedents with null pronouns and overt-pronoun-*ga* in her sentence-interpretation experiments. An example sentence from Nagano (2014) can be seen below.

(23)

Takeshi-san-wa/ga biiru-o nondeiru aida, Ø/kare-ga keshiki-o nagameteimasu.
 Takeshi-TOP/NOM beer-ACC drinking while Ø/he-NOM scenery-ACC view
 'While Takeshi drinks beer, he views the scenery.'

(Nagano, 2014)

Nagano found that, while participants tended to interpret the null pronouns as referring to subject antecedents (*Takeshi*), they tended to interpret overt pronouns as referring to non-subject antecedents (someone other than *Takeshi*).²³ In spite of some flux in the usage of Japanese overt pronouns, a consensus among Ueno and Kehler (2009), Okuma (2010) and Nagano (2014) is that null pronouns prefer subject antecedents consistently and much more strongly than overt pronouns do.

Empathy and Antecedent Accessibility in Japanese

Walker, Iida and Cote (1994) maintain that Japanese antecedents' accessibility is also enhanced by empathy. That is, an antecedent that is an empathy locus is more accessible than an antecedent that is a non-empathy locus. The notion of empathy was proposed by Kuno (1976) in order to present the speakers' position in describing a situation. In general, empathy indicates the point of view or perspective of a person that a

²² Okuma could not use overt-pronoun-*wa* because she used the anaphor in embedded clauses, which do not allow the use of *wa*.

²³ Nagano does not explain the difference between *wa* and *ga* that marked the antecedents.

speaker empathizes with. Refer to the following sentences, both of which describe a situation where a father Taro hugged his son Saburo.

(24) a. Taro hugged his son (= Saburo).

b. Saburo's father (= Taro) hugged him. (Walker et al., 1994, p. 208)

In (24a), the sentence is describing the father using his name *Taro*, and describing the son using the term *his son*. These ways to address these two people foreground *Taro* and background *Saburo*. In other words, the speaker of this sentence is empathizing with *Taro*'s point of view, so *Taro* is the empathy locus. In contrast, (24b) is describing the son using his name *Saburo*, and describing the father using the term *father*. These foreground *Saburo* and background *Taro*. The speaker is empathizing with *Saburo*, so *Saburo* is the empathy locus. Walker et al. suggest that *Taro* is more accessible than *Saburo* in (24a), and *Saburo* is more accessible than *Taro* in (24b).

Since empathy is a non-linguistic notion, it does not have an independent NP category and always overlaps the other categories (including the information-structural category of the topic, focus, etc. and grammatical categories of the subject, object, etc.). According to Kuno (1987), speakers tend to empathize with the subject more than the non-subject and empathize with the topic more than non-topic. In other words, when a speaker says a sentence such as "John hit Mary", the speaker tends to be empathizing with *John* (although it is possible that this sentence is a neutral description that does not overtly express any specific empathy locus). The empathy locus associating with the subject can be adjusted by passivization. For instance, the sentences "John hit Mary" and "Mary was hit by John" describe the same situation, but the speaker of the sentences empathizes with the subject, namely *John* in the former sentence and *Mary* in the latter

sentence.²⁴

Instead of objectivity, Japanese speakers tend to focus on their empathy loci on a certain character/person in narratives (Nakahama, 2011). This is distinct from English, which focuses more on fact and objectively describes a situation. Uehara (1996) shows this difference in the English sentences taken from *The Last Leaf* by O. Henry (Poter, 1948) and the equivalent Japanese sentences translated by Yasuo Ohkubo (1969).

(25) a. Original English sentence: Then I saw a big lady standing there.

Japanese translation: Futotta obasan-ga tatteita.

‘A fat woman was standing.’

b. Original English sentence: ...she found Johnsy with dull wide-open eyes
staring at the drawn green shade.

Japanese translation: ...Jonsi-wa, seiki-no nai me-o ookiku mihiraite,
orosareteiru midori-iro-no syeedo-o, zitto
mitumeteita. ‘Johnsy was staring at the drawn
green shade with dull wide-open eyes.’

(Uehara, 1996, pp. 212-213)

In the Japanese translations, any equivalent pronoun to *I* or *she* is non-existent. While the English sentences are described from objective viewpoints, the Japanese sentences are described from subjective viewpoints of the characters in the novel (i.e., *I*, *she*). In other words, the English sentences are described from a bird’s-eye view, which captures two

²⁴ The likelihood of an NP being the empathy locus can be also dependent on various factors including animacy, person, social situation and topicality (Kuno, 1987; Oshima, 2007).

characters (e.g., *I* and *the big lady*; *she* and *Johnsy*), and the Japanese sentences are described from the eyes of *I* or *she*, which capture only *the big lady* or *Johnsy*. In other words, the teller of this story in Japanese translation is empathizing with *I* or *she*. Iwasaki (1993) and Kanaya (2004) maintain that, unlike English, the subjective perspective plays a prominent role in Japanese. Iwasaki distinguishes these subjective and objective perspectives by calling them S(elf)-perspective (i.e., “subjective experiencing self”) and O(ther)-perspective (i.e., “objective observing self”), respectively (Iwasaki, 1993, p. 18). Similar instances of S-perspective can be seen in English, often with the uses of perception verbs such as *see*, *hear*, *look*, *sound*, and so forth.

(26) Dan went to a party yesterday. He saw his high school friend Jim. He looked awfully pale. (Kameyama, 1986, p. 205)

The third sentence in (26) is described from *Dan*’s perspective (i.e., through the eyes of *Dan*, *Jim* looked awfully pale) although the sentence does not include *Dan* or a pronoun that refers to *Dan*. However, Japanese employs the S-perspective more often than English, and it is implicitly encoded from whose perspective the sentences are described (Iwasaki, 1993, p.80).

This tendency to empathize with a certain character or person in a sentence is also found in some certain Japanese words. For example, some verb phrases obligatorily express the empathy locus, regardless of the grammatical status of the NPs. An example is in the verbs describing an event involving giving and receiving (Kuno, 1987). In Japanese, there is no way to objectively describe a giving and receiving situation, according to Kuno and Kaburaki (1977). Two different verbs for *give* are utilized in order to explicitly express the empathy loci or the points of views, from which the sentences

are described. More concretely, the two verbs for ‘give’, namely *ageru* and *kureru*, are differentiated in order to indicate whether the sentence is described from the giver’s viewpoint or the receiver’s viewpoint, respectively. In the following example sentences, the giver in (27a) and the receiver in (27b) are expressed by *watasi* ‘I’, which is the most obvious term that expresses the empathy locus.

- (27) a. *Watasi-ga* *Taro-ni* *hon-o* *ageta*/**kureta*.
 I-NOM Taro-DAT book-ACC gave
 ‘I gave Taro a book.’
- b. *Taro-ga* *watasi-ni* *hon-o* *kureta*/**ageta*.²⁵
 Taro-NOM I-DAT book-ACC gave
 ‘Taro gave me a book.’

Both *ageta* and *kureta* mean *gave*, but *ageta* is used when the sentence is described from the giver’s perspective. Thus, *ageta* has to be used in sentence (27a) because the giver, ‘I’, is the empathy locus. On the other hand, *kureta* is used when the sentence is described from the receiver’s perspective. Thus, *kureta* has to be used in sentence (27b) because the receiver, ‘I’, is the empathy locus.

Even when the giver or receiver is not the speaker himself/herself, a speaker can differentiate *ageta* and *kureta* in his/her utterance depending on whether the speaker is empathizing with the giver or with the receiver. Look at the minimally paired example sentences below, both of which mean, ‘Taro gave Ziro a book’.

²⁵ *Ageta* and *kureta* in (27) are the past tenses of *ageru* ‘give’ and *kureru* ‘give’, respectively.

- (28) a. Taro-ga Ziro-ni hon-o ageta.
 Taro-NOM Ziro-DAT book-ACC gave
 ‘Taro gave Ziro a book.’
- b. Taro-ga Ziro-ni hon-o kureta.
 Taro-NOM Ziro-DAT book-ACC gave
 ‘Taro gave Ziro a book.’

Since *ageta* is used when the speaker is empathizing with the giver’s perspective, sentence (28a) with *ageta* indicates that giver (i.e., *Taro*) is the empathy locus. In other words, the speaker of this sentence is empathizing with *Taro*. On the other hand, *kureta* should be used when the speaker is empathizing with the receiver, so (28b) with *kureta* indicates that the receiver (i.e., *Ziro*) is the empathy locus. In other words, the speaker is empathizing with *Ziro*. Typically, *kureta* is used when a speaker describes someone giving a gift to him/her, someone giving a present to his/her son or daughter, or someone in another company giving a souvenir to his/her co-worker.

In addition, the words, *ageru* and *kureru* can be used as auxiliary verbs that affix other verbs in order to mean, ‘did for (someone)’ or ‘give (someone) a favor of –ing’. Examples are shown below.

- (29)
- a. Taro-ga Ziro-ni hon-o katte-ageta.
 Taro-NOM Ziro-DAT book-ACC gave a favor of buying
 ‘Taro bought a book for Ziro (Taro gave a favor of buying a book for Ziro).’
- b. Taro-ga Ziro-ni hon-o katte-kureta.
 Taro-NOM Ziro-DAT book-ACC gave a favor of buying
 ‘Taro bought a book for Ziro (Taro gave a favor of buying a book for Ziro).’

In the sentences above, the verbs *ageta* and *kureta* ‘gave’ affix another verb *kau* ‘buy’ (in an inflected form *katte*), so they together construct *katte-ageta/katte-kureta* ‘gave a favor of buying’. These combinations of *ageru/kureru* ‘give’ and another verb in the inflected

form of *-te* show the same empathy hierarchy as *ageru/kureru* (Kuno, 1987). In other words, *-te ageru* indicates that giver is the empathy locus, and *-te kureru* indicates that receiver is the empathy locus. *Taro* is the empathy locus in sentence (29a) because *katte-ageta* is used and *Taro* is the giver; *Ziro* is the empathy locus in (29b) because *katte-kureta* is used and *Ziro* is the receiver. Yanagimachi (2000) reports that, in his observation of Japanese speakers' storytelling, they used these auxiliary verbs of giving and receiving in order to set their empathy on the protagonist of the story. In other words, they used *-te ageru* when the protagonist was doing a favor, and they used *-te kureru* when the protagonist was receiving a favor. This strategy helps their narrations remain coherent and comprehensible to the listeners (Nakahama, 2011).

Another example that expresses different empathy loci is *iku* 'go' and *kuru* 'come' (Kawakami, 1996; Kuno, 1978; Koga, 2014). When one says that someone "goes" to somewhere, the speaker is empathizing with the person who is going, but if he/she says someone "comes" to somewhere, the speaker is empathizing with the destination.

Examples are shown below.

- (30) a. Taro-ga Ziro-no uti-ni itta.
 Taro-NOM Ziro-GEN house-DIR went
 'Taro went to Ziro's house.'
- b. Taro-ga Ziro-no uti-ni kita.²⁶
 Taro-NOM Ziro-GEN house-DIR came
 'Taro came to Ziro's house.'

The empathy of the speaker of this sentence is placed on *Taro* in (30a), while the speaker's empathy is placed on *Ziro* in (30b).²⁷

²⁶ *Itta* and *kita* in (30) are the past tenses of *iku* 'go' and *kuru* 'come', respectively.

The verbs *iku* ‘go’ and *kuru* ‘come’ combined with *motu* ‘hold’ make other verbs *motte-iku* and *motte-kuru*, both of which mean ‘bring’. When a sentence describes that someone brings something to someone, the use of *motte-iku* expresses that the empathy locus is the person who brings something, and the use of *motte-kuru* expresses that the empathy locus is the person who receives the thing. Refer to example sentences below.

- (31) a. Taro-ga Ziro-ni purezento-o motte-itta.
 Taro-NOM Ziro-DAT present-ACC brought
 ‘Taro brought a present to Ziro.’
- b. Taro-ga Ziro-ni purezento-o motte-kita.²⁸
 Taro-NOM Ziro-DAT present-ACC brought
 ‘Taro brought a present to Ziro.’

In sentence (31a) with *motte-itta*, the speaker empathizes with *Taro*, so *Taro* is the empathy locus. The situation could be that the speaker is *Taro*’s close friend, and he/she is describing that *Taro* is bringing a present to someone (i.e., *Ziro*). In contrast, in sentence (31b) with *motte-kita*, the speaker empathizes with *Ziro*, so *Ziro* is the empathy

²⁷ Note that the usages of *go* and *come* are slightly different between Japanese and English. While the usages are dependent on whom the speaker of the sentence empathizes with in Japanese, they are dependent on whom the listener empathizes with in English. For example, in English, John’s brother can say to Mary’s sister, “John is coming to Mary’s house” although John’s brother is empathizing with John. This is because the listener, Mary’s sister is empathizing with Mary.

²⁸ *Motte-itta* and *motte-kita* in (31) are the past tenses of *motte-iku* ‘bring’ and *motte-kuru* ‘bring’, respectively.

locus. The situation could be that the speaker is *Ziro*'s close friend, and he/she is describing that someone (i.e., *Taro*) is bringing a present to *Ziro*.

Another instance shows that *iku* 'go' and *kuru* 'come' are combined with a verb *kau* 'buy'. These combinations derive *katte-iku* and *katte-kuru*, both of which mean 'buy (for someone)'. These words are used to describe a situation where a person buys something (typically as a gift) for another person. *Katte-iku* expresses that the empathy locus is the person who buys something, and *katte-kuru* expresses that the empathy locus is the person who receives the thing. Example sentences are shown below.

- (32) a. Taro-ga Ziro-ni purezento-o katte-itta.
 Taro-NOM Ziro-DAT present-ACC bought
 'Taro bought a present for Ziro.'
- b. Taro-ga Ziro-ni purezento-o katte-kita.²⁹
 Taro-NOM Ziro-DAT present-ACC bought
 'Taro bought a present for Ziro.'

The sentence (32a) with *katte-itta* indicates that the speaker of this sentence empathizes with *Taro*, and the sentence (32b) with *katte-kita* indicates that the speaker empathizes with *Ziro*.

If the empathy-locus antecedents are more accessible than non-empathy-locus antecedents as Walker, Iida and Cote (1994) suggest, it is implied that empathy loci are more salient than non-empathy loci. Therefore, empathy-locus antecedents would require an anaphor that are immediately interpreted as an anaphor (according to the DPT) or an anaphor that carries less semantic information (according to the ILH), and non-empathy-

²⁹ *Katte-itta* and *katte-kita* in (32) are the past tenses of *katte-iku* 'buy' and *katte-kuru* 'buy', respectively.

locus antecedents would accept an anaphor that are not immediately interpreted as an anaphor or an anaphor that carries rich information. Accordingly, the RNP and OPP would be elicited with empathy-locus antecedents and non-empathy-locus antecedents. In other words, the second sentence in a Japanese discourse such as “Taro_i gave (*ageta*) Hanako a present. Ø_i knew that today is Hanako’s birthday.” would be read faster than that in a discourse such as “Taro_i gave (*kureta*) Hanako a present. Ø_i knew that today is Hanako’s birthday.” In the former discourse, *ageta* indicates *Taro* is the empathy locus, so *Taro* is a salient antecedent. In the latter discourse, *kureta* indicates *Hanako* is the empathy locus, so *Taro* is a less salient antecedent.

The above prediction regarding the RNP and OPP with empathy-locus antecedents and non-empathy-locus antecedents agrees with Clancy’s (1980) and Nakahama’s (2011) finding that, in Japanese, a null pronoun is most often used when referring to a main character, which is most likely the speakers’ empathy locus. This prediction regarding the OPP could be also supported by our earlier suggestion that Japanese overt pronouns are anti-logophoric epithet-pronouns that tend to avoid referring to the person from whose perspective the sentence is described. According to Sells (1987), the notion of logophoricity roughly corresponds to empathy. In the present OPP experiments, Japanese overt pronouns may not refer to empathy-locus antecedents, if these antecedents are subject to anti-logophoricity.

This chapter discussed possibly salient or highly accessible antecedents such as subject (as a grammatical category), topic (as an information-structural category) and empathy locus (as a non-linguistic category), all of which were tested in the experiments of the current study.

Chapter 5. Current Study

The current study examined the effects of the RNP and OPP in Japanese with different combinations of anaphors and antecedents. Following the work of Gordon, Grosz and Gilliom (1993) and Gelormini-Lezama and Almor (2011), six self-paced, sentence-by-sentence E-Prime reading experiments were conducted. These experiments measured the reading times of sentences that included anaphors that referred to antecedents in preceding sentences. I chose the sentence-by-sentence paradigm based on the use of this paradigm in previous studies and given evidence that word-by-word reading paradigms do not elicit the RNP (Nair & Almor, 2006). The data was entirely collected from native, Japanese-speaking university students residing in Japan. The structures of the six experiments of this study are shown in Table 5.1 below. Experiments 1, 3 and 5 examined the RNP, so the anaphors used were null pronouns and repeated names. Experiments 2, 4 and 6 examined the OPP, so the anaphors were null pronouns and overt pronouns. All the anaphors used were grammatical subjects. Also, Experiments 1 and 2 used subject antecedents and non-subject (object) antecedents as possibly more salient antecedents and less salient antecedents, respectively. Experiments 3 and 4 used topic antecedents and non-topic antecedents (both were subjects) as possibly more salient antecedents and less salient antecedents, respectively. Experiments 5 and 6 used empathy-locus antecedents and non-empathy-locus antecedents (both were subjects) as possibly more salient antecedents and less salient antecedents, respectively.

Table 5.1

Antecedents and Anaphors in Experiments 1-6

	<u>Antecedent</u>	<u>Anaphor (Subject)</u>
Experiment 1 (testing RNP)	Subject vs.	Null pronoun vs.
	<u>Non-subject (Object)</u>	<u>Repeated-name-ga/wa</u>
Experiment 2 (testing OPP)	Subject vs.	Null pronoun vs.
	<u>Non-subject (Object)</u>	<u>Overt-pronoun-ga/wa</u>
Experiment 3 (testing RNP)	Topic (Subject) vs.	Null pronoun vs.
	<u>Non-topic (Subject)</u>	<u>Repeated-name-ga/wa</u>
Experiment 4 (testing OPP)	Topic (Subject) vs.	Null pronoun vs.
	<u>Non-topic (Subject)</u>	<u>Overt-pronoun-ga/wa</u>
Experiment 5 (testing RNP)	Empathy-locus (Subject) vs.	Null pronoun vs.
	<u>Non-empathy-locus (Subject)</u>	<u>Repeated-name-ga/wa</u>
Experiment 6 (testing OPP)	Empathy-locus (Subject) vs.	Null pronoun vs.
	<u>Non-empathy-locus (Subject)</u>	<u>Overt-pronoun-ga/wa</u>

For data collection, all the participants were residents of Japan, native Japanese speakers, and were undergraduate or graduate students in Japanese universities. In order to ensure that English or any other second language did not influence their linguistic instincts, no one was allowed to participate if he/she had stayed in a foreign country for one continuous month or more.

For the data analysis throughout all experiments I measured the reading times of

the second sentences that included anaphors. Reading times from trials in which the participant answered the comprehension question incorrectly were removed prior to the analyses. In addition, I utilized a conservative method for outlier removals that only removed extreme reading times below 300 ms or greater than 6000 ms, as they clearly reflected either equipment or subject error. In addition, the data of the participants whose comprehension accuracy rate was below 80% were removed from all the experiments.

I analyzed log-transformed reaction times in R version 3.2.2 using mixed effects models with R package lme4.³⁰ I compared different models to estimate the significance of each term starting with the maximal model containing the antecedent and anaphor and their interaction as fixed-effect conditions. I first tried to eliminate the interaction term, and if this elimination did not result in a significant loss of model fit, I then attempted to remove each of the individual factors (Baayen, 2008). As recommended by Barr, Levy, Scheepers and Tily (2013), I included the maximal structure of by-participant and by-item random intercepts and slopes that allowed the models to converge. I report the coefficients in the final models. Throughout the six experiments, I used the same models for the data analysis.

³⁰ Log-transformation makes data more interpretable when the data is highly skewed (Lane, D. M., <http://onlinestatbook.com/2/transformations/log.html>).

Chapter 6. Experiments 1 and 2

Predictions

Experiments 1 and 2 were standard RNP/OPP experiments with subject antecedents and object antecedents, but also contained overt anaphors which are differentiated between the topic and non-topic by using *wa* and *ga*. These experiments investigated the following issues:

- (i) RNP: Participants may read sentences with null pronouns faster than sentences with repeated-name anaphors when the antecedents are salient (i.e., subject). However, the reading-time difference might significantly decrease when the antecedents are not salient (i.e., object).
- (ii) OPP: Participants may read sentences with null pronouns faster than sentences with overt pronouns when the antecedents are salient (i.e., subject). However, the reading-time difference might significantly decrease when the antecedents are not salient (i.e., object).
- (iii) Topic: Participants may process sentences with the overt anaphors differently depending on whether they are the topic (i.e., repeated-name topic-*wa* and overt-pronoun topic-*wa*) or the non-topic (i.e., repeated-name non-topic-*ga* and overt-pronoun non-topic-*ga*).

Participants

Forty-two undergraduate students from Mie University in Japan participated in Experiment 1. They were all native Japanese speakers and between 18 and 22 years of age ($M = 19.833$). Thirty-eight graduate or undergraduate students from the University of

Tokyo participated in Experiment 2. They were all native Japanese speakers and between 19 and 30 years of age ($M = 19.974$). Each session lasted approximately 20 minutes.

Materials

Two-sentence discourses were given throughout a session: the first sentences included antecedents, and the second sentences included anaphors. In both Experiments 1 and 2, antecedents were the subject and object, which were referred to by three different types of anaphors. In Experiment 1, examining RNP, anaphors were repeated names with *ga* (i.e., repeated-name non-topic), repeated names with *wa* (i.e., repeated-name topic), and null pronouns (i.e., null topic). In Experiment 2, examining OPP, anaphors were overt pronouns with *ga* (i.e., overt-pronoun non-topic) overt pronouns with *wa* (i.e., overt-pronoun topic) and null pronouns (i.e., null topic). All the anaphors were grammatical subjects, which are most likely to elicit the RNP and OPP, according to previous studies. Six items were provided for each condition, which creates 36 experimental items. These were mixed among 48 distractors. Sample items are shown in Tables 6.1 and 6.2 below.

Although we could not help that the sentences with null pronouns were shorter than the others, the lengths of the second sentences were kept approximately the same by aligning the number of letters and morae. Namely, for the sentences with null pronouns in Experiments 1 and 2, the number of letters and morae ranged from 15 to 18 and 18 to 22, respectively. For the sentences with overt pronouns and repeated names, they ranged from 17 to 20 and from 22 to 26, respectively.

Table 6.1

Conditions and Example Items for Experiment 1 (RNP)

Conditions:	Antecedent	Anaphor
	Subject	Repeated name <i>-ga</i>
	Object	Repeated name <i>-wa</i>
		Null pronoun
<hr/>		
<u>Example items:</u>		
<hr/>		
First sentence with Antecedent		
(i) Subject		
e.g.	<u>Taku-ga</u> Kazuko-o arubaito tosite yatotta. Taku-NOM Kazuko-ACC part-timer as hired 'Taku hired Kazuko as a part-timer.'	
(ii) Object		
e.g.	Kazuko-ga <u>Taku-o</u> arubaito tosite yatotta. Kazuko-NOM Taku-ACC part-timer as hired 'Kazuko hired Taku as a part-timer.'	
<hr/>		
Second sentence with Anaphor		
(i) Repeated name <i>-ga</i>		
e.g.	Taku-ga mainiti Kazuko-to issmoni hataraita. Taku-NOM everyday Kazuko-with together worked 'Taku worked with Kazuko everyday.'	
(ii) Repeated name <i>-wa</i>		
e.g.	Taku-wa mainiti Kazuko-to issmoni hataraita. Taku-TOP everyday Kazuko-with together worked 'Taku worked with Kazuko everyday.'	
(iii) Null pronoun		
e.g.	Ø mainiti Kazuko-to issmoni hataraita. everyday Kazuko-with together worked '(he) worked with Kazuko everyday.'	
<hr/>		

Table 6.2

Conditions and Example Items for Experiment 2 (OPP)

Conditions:	Antecedent	Anaphor
	Subject	Overt pronoun <i>-ga</i>
	Object	Overt pronoun <i>-wa</i>
		Null pronoun

Example items:

First sentence with Antecedent

(i) Subject

e.g. Taku-ga Kazuko-o arubaito tosite yatotta.
 Taku-NOM Kazuko-ACC part-timer as hired
 ‘Taku hired Kazuko as a part-timer.’

(ii) Object

e.g. Kazuko-ga Taku-o arubaito tosite yatotta.
 Kazuko-NOM Taku-ACC part-timer as hired
 ‘Kazuko hired Taku as a part-timer.’

Second sentence with Anaphor

(i) Overt pronoun *-ga*

e.g. Kare-ga mainiti Kazuko-to isshoni hataraita.
 he-NOM everyday Kazuko-with together worked
 ‘He worked with Kazuko everyday.’

(ii) Overt pronoun *-wa*

e.g. Kare-wa mainiti Kazuko-to isshoni hataraita.
 he-TOP everyday Kazuko-with together worked
 ‘He worked with Kazuko everyday.’

(iii) Null pronoun

e.g. Ø mainiti Kazuko-to isshoni hataraita.
 everyday Kazuko-with together worked
 ‘(he) worked with Kazuko everyday.’

In order to make Experiments 1 and 2 as uniform as possible, the repeated-name anaphor versions were the same length as the overt-pronoun versions. The overt pronouns, ‘he’ and ‘she’, in Japanese are *kare* (彼) and *kanojo* (彼女), written in one character and two characters, respectively. The names used in the items had the same number of letters as the overt pronouns (e.g., *Taku* or 拓 as a male name and *Kazuko* or 和子 as a female name). All of the experimental discourses included two persons’ names, which were one male and one female (e.g., *Taku* and *Kazuko*). The number of sentences with male and female anaphors was balanced across the items and in the individual lists shown to each participant. Names used in the experiments were all familiar ones to native Japanese speakers. These names’ familiarities were solicited via questionnaire from twelve native Japanese speakers. The ages of these Japanese speakers for this familiarity check (i.e., mean age was 22.667, ranging from 19 to 27) were similar to those of the participants of Experiments 1 through 6 (i.e., mean age was 19.256, ranging from 18 to 30). In the familiarity-check questionnaire, the twelve native Japanese speakers were provided with a list of 120 common Japanese names (60 male and 60 female names), and they rated the commonness of them with a Likert scale consisting of five options: 1 is “very common”, 2 is “common”, 3 is “neither common nor uncommon”, 4 is “uncommon”, and 5 is “very uncommon”.³¹ The top 36 male names and 36 female names, whose average ratings were

³¹ The 120 names in the list were taken from the popular Japanese first-name ranking websites (<http://www.namaejiten.com>; http://coreblog.org/naming/fp/ranking/year_2010.html; <http://tamahiyo.jp/namae/2012/name-ranking.html>; http://www1.odn.ne.jp/haru/data-other/name-00_m.html).

below 2.5, were used in the experimental discourses. Average ratings of the names used were 1.887 (ranging from 1.083 to 2.417) for male names and 1.838 (ranging from 1.167 to 2.417) for female names. Also, regarding the orthographies of words used in the experimental discourses, since the uses of kanji script or kana script for written words affect Japanese speakers' processing of the words (Kashiwagi & Nakayama, 2008), each word used was written in a more common script, e.g., けんか 'fight' rather than 喧嘩 'fight'; お辞儀 'bow' rather than おじぎ 'bow'. The more common script was determined by Kotonoha Shonagon corpus. By looking up the same words in kana scripts, e.g., けんか 'fight', and kanji scripts, e.g., 喧嘩 'fight', I used the script that were found to have a larger number of instances in the corpus. In addition, because the use of the nominative-postposition *ga* is markedly unnatural in copula sentences (Mio, 1948/2003), all the experimental sentences were non-copula sentences. Each discourse was followed by a yes/no comprehension question (e.g., *Mondai: Taku-ga Kazuko-o yatoimashita ka.* 'Question: Did Taro hire Kazuko?').

A concern regarding the experimental item is that the second sentence might need to start with a connective phrase such as *soshite* 'and' or *sorekara* 'then', which explicitly maintains the cohesion between the first and second sentences. Compared to English, Japanese utilizes connectives more frequently (Nishihara, 1990). The second sentence with connectives may be more smoothly read continuously after the first sentence, and thus there is a possibility that readers might feel the discourse somewhat unnatural without a connective phrase. However, using connective phrases might produce a continuing discourse effect, which would give a bias favoring the subject or object referents. It is well known that connectives could lead readers to build a reference to

antecedents in a particular grammatical position or theta role (Kaiser, Li & Holsinger, 2011). Because the antecedent bias would greatly interfere with our purpose in the present study, we avoided using connectives.

Procedure

The basic experimental design followed that of Gelormini-Lezama and Almor (2011). Each of the 36 experimental discourses that the participants read was presented in one of the six conditions, in a Latin-Square manner. Thus, six of the discourses for each condition were provided to each participant. Also, prior to the actual experiment, a practice block consisting of four discourses was provided in order to familiarize the participants with their experimental task.

In both experimental and practice sessions, participants read the discourses sentence-by-sentence in a self-paced fashion. Each trial began with a “+” fixation mark, which was followed by the first sentence, the second sentence and then the comprehension question. Participants used the space bar to advance to the next screen and the 1 and 2 number keys to answer yes and no, respectively, to the comprehension questions. The experimental and distractor discourses were mixed and presented in a different random order for each participant.

Results: Experiment 1

I measured the reading times for the second sentences that included anaphors. For the data from the 42 participants, reading times from trials in which the participant answered the comprehension question incorrectly were removed, affecting 5.09% of the data. In addition, after examining the data, a criterion of removing extreme reading times below 300 ms or greater than 6000 ms was chosen, affecting 1.25%. The reading times of

each type of anaphor are shown in Figure 6.1 below. Error bars in the figure represent the standard error of the mean.

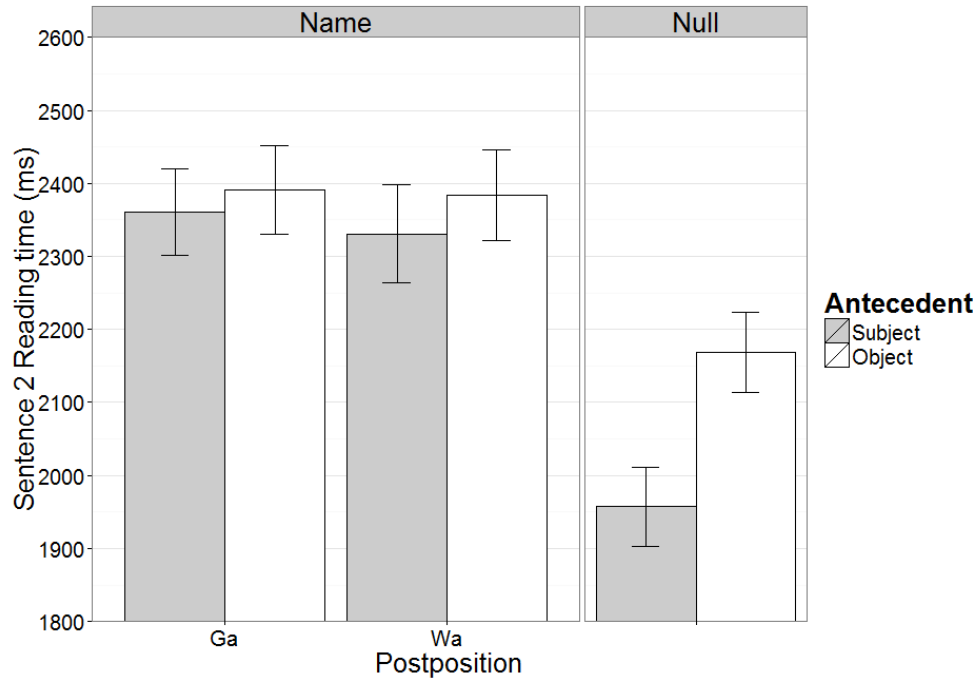


Figure 6.1. Reading times of the second sentences with anaphors in Experiment 1

Using the model of Gelormini-Lezama and Almor (2011), I, therefore, considered the interaction between anaphor form (repeated names vs. null pronouns, overt pronouns vs. null pronouns) and antecedent (subject vs. object) as a reflection of the RNP or OPP. Indeed, our omnibus analysis revealed a marginally significant interaction between antecedent (subject vs. object) and anaphor (repeated-name-*ga*, repeated-name-*wa*, and null pronoun) ($\chi^2(2) = 4.84, p = .089$). Table 6.3 shows the coefficients of the fixed terms in the final model. Importantly, the coefficients of the interaction terms in this analysis

were significant with the repeated-name non-topic-subject-*ga*, which suggests a RNP, but the repeated-name topic-subject-*wa* did not show the interaction.

Table 6.3

Omnibus RNP Analyses for Experiment 1

	β	SE	T	P
Intercept	7.490	0.040	185.445	< .001**
Antecedent: Object	0.111	0.032	3.490	< .001**
Anaphor: Repeated-name-Ga	0.196	0.032	6.186	< .001**
Anaphor: Repeated-name-Wa	0.178	0.032	5.621	< .001**
Antecedent: Object \times Anaphor: Repeated-name-Ga	-0.096	0.045	-2.147	= .032*
Antecedent: Object \times Anaphor: Repeated-name-Wa	-0.067	0.045	-1.487	= .137

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

To further confirm these findings, I conducted a series of analyses aiming to detect the RNP for each full reference form separately. The analyses contrasted repeated-name-*ga* vs. null pronouns and repeated-name-*wa* vs. null pronouns. The results are

summarized in Tables 6.4 and 6.5 below. In agreement with the omnibus analysis above, these analyses revealed that the interaction between anaphors and antecedents, when anaphors included repeated-name non-topic-*ga*, i.e., $\chi^2(1) = 4.88$, $p = .03$, but not repeated-name topic-*wa*, i.e., $\chi^2(1) = 2.06$, $p = .15$.

Table 6.4

Separate RNP Analysis for Experiment 1: Null Pronoun vs. Repeated-name Non-topic-ga

	β	SE	T	P
Intercept	7.490	0.042	177.303	< .001**
Antecedent: Object	0.112	0.031	3.581	< .001**
Anaphor: Repeated-name-Ga	0.196	0.031	6.287	< .001**
Antecedent: Object \times Anaphor: Repeated-name-Ga	-0.098	0.044	-2.213	= .027*

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 6.5

Separate RNP Analysis for Experiment 1: Null Pronoun vs. Repeated-name Topic-wa

	β	SE	T	P
Intercept	7.490	0.039	189.889	< .001**
Antecedent: Object	0.111	0.032	3.450	< .001**
Anaphor: Repeated-name-Wa	0.177	0.032	5.534	< .001**
Antecedent: Object \times Anaphor: Repeated-name-Wa	-0.066	0.048	-1.438	= .151

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The interaction in repeated-name non-topic-*ga* indicated smaller reading-time differences between repeated-name non-topic-*ga* and the null pronoun in the object-antecedent condition than in the subject-antecedent condition consistent with the RNP (Gordon, Grosz & Gilliom, 1993). In contrast, the reading-time differences between repeated-name topic-*wa* and null pronoun were not significantly different between subject- and object-antecedent conditions.

I also conducted simple comparisons contrasting the reading time of each anaphor form between the subject- and object-antecedent conditions.

Table 6.6

Simple Subject vs. Object Antecedent Effects for Each Anaphor Form in Experiment 1

	β	SE	T	P
Null pronoun	0.111	0.032	3.490	< .001**
Repeated-name-Ga	0.010	0.031	0.330	= .741
Repeated-name-Wa	0.059	0.032	1.850	= .065 .

Note. Factors were coded with dummy coding. The subject antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The coefficients corresponding to these simple comparisons indicate that null pronouns were the only anaphors significantly read faster for subject antecedents than object antecedents. Also, repeated-name-*wa* were read faster for subject antecedents than object antecedents at a marginal significance. However, repeated-name-*ga* did not show a difference.

Results: Experiment 2

For the dataset involving 38 participants, reading times from trials to which the participants answered the comprehension question incorrectly were removed, which affected 3.07% of the data. In addition, outlier data, defined as reading times below 300

ms or greater than 6000 ms were removed, which affected 0.83%. The reading times of each type of anaphor are shown in Figure 6.2 below. Error bars in the figure represent the standard error of the mean³²

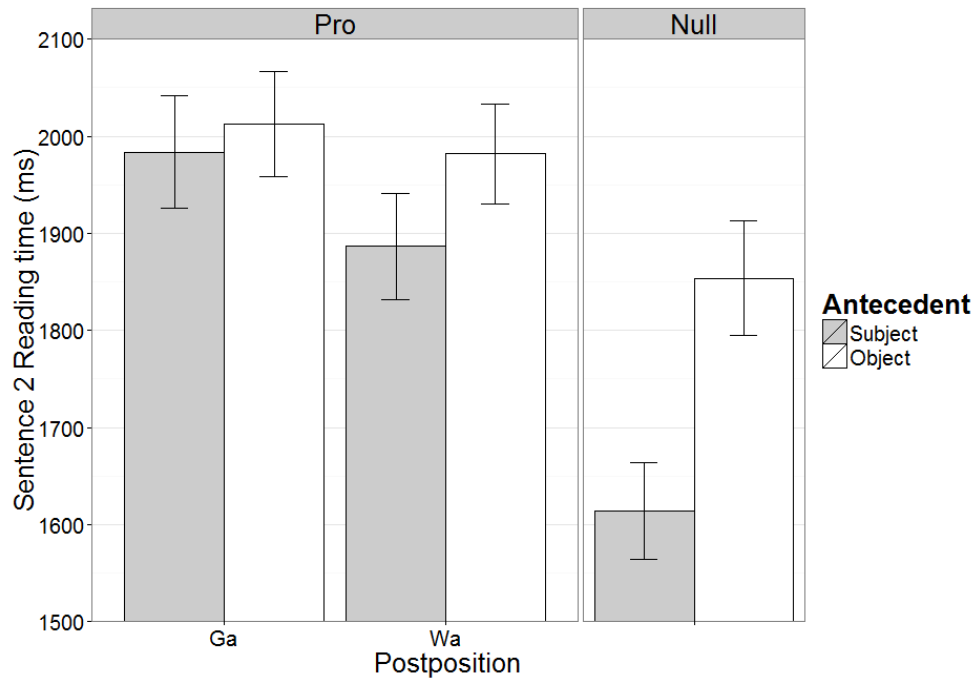


Figure 6.2. Reading times of the second sentences with anaphors in Experiment 2

³² The raw reading times of Experiment 2 were faster than Experiment 1, which might reflect a baseline difference between the participants in Experiments 1 (students of Mie University) and 2 (students of the University of Tokyo). The University of Tokyo is academically ranked higher than Mie University, (<http://daigaku.ehoh.net/hensachi/index.html>), so the students of the University of Tokyo could have processed information faster than the students of Mie University.

Our omnibus analysis found a marginally significant interaction between antecedent (subject vs. object) and anaphor (overt-pronoun-*ga*, overt-pronoun-*wa*, and null pronoun) ($\chi^2(2) = 5.55, p = .062$). Table 6.7 shows the coefficients of the fixed terms in the final model. Similar to Experiment 1, the results showed that the coefficients of the interaction terms were significant with the overt-pronoun non-topic-subject-*ga*. This suggests the OPP, but the overt-pronoun topic-subject-*wa* did not show this interaction.

Table 6.7

Omnibus OPP Analyses for Experiment 2

	β	SE	T	P
Intercept	7.299	0.045	162.537	< .001**
Antecedent: Object	0.121	0.032	3.792	< .001**
Anaphor: Overt-pronoun-Ga	0.209	0.032	6.539	< .001**
Anaphor: Overt-pronoun-Wa	0.150	0.032	4.713	< .001**
Antecedent: Object \times Anaphor: Overt-pronoun-Ga	-0.106	0.045	-2.358	= .019*
Antecedent: Object \times Anaphor: Overt-pronoun-Wa	-0.052	0.045	-1.151	= .250

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Like Experiment 1, a series of analyses aiming to detect the OPP for each reference form was separately conducted. The analyses compared overt-pronoun-*ga* vs. null pronouns and overt-pronoun-*wa* vs. null pronouns. The results are summarized in Tables 6.8 and 6.9 below. These analyses revealed an interaction in overt-pronoun non-topic-*ga*, i.e., $\chi^2(1) = 5.17$, $p = .02$, but not in overt-pronoun topic-*wa*, i.e., $\chi^2(1) = 1.30$, $p = .25$.

Table 6.8

Separate OPP Analysis for Experiment 2: Null Pronoun vs. Overt-pronoun Non-topic-ga

	β	SE	T	P
Intercept	7.298	0.045	163.707	< .001**
Antecedent: Object	0.120	0.033	3.666	< .001**
Anaphor: Overt-pronoun-Ga	0.208	0.033	6.328	< .001**
Antecedent: Object \times Anaphor: Overt-pronoun-Ga	0.047	0.047	-2.278	= .023*

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 6.9

Separate OPP Analysis for Experiment 2: Null Pronoun vs. Overt-pronoun Topic-wa

	β	SE	T	P
Intercept	7.301	0.046	157.819	< .001**
Antecedent: Object	0.121	0.032	3.834	< .001**
Anaphor: Overt-pronoun-Wa	0.145	0.031	4.621	< .001**
Antecedent: Object \times Anaphor: Overt-pronoun-Wa	-0.051	0.045	-1.143	= .254

Note. Factors were coded with dummy coding. The null pronoun and subject antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The interaction in overt-pronoun non-topic-*ga* indicated the OPP (Gelormini-Lezama & Almor, 2011); smaller reading-time differences between overt-pronoun non-topic-*ga* and the null pronoun in the object-antecedent conditions than the reading-time differences present in the subject-antecedent conditions. On the other hand, the reading-time differences between the overt-pronoun topic-*wa* and null pronoun were not significantly different between subject- and object-antecedent conditions.

Like Experiment 1, simple comparisons, which contrast the reading times of each anaphors between the subject- and object-antecedent conditions, were conducted.

Table 6.10

Simple Subject vs. Object Antecedent Effects for Each Anaphor Form in Experiment 2

	β	SE	T	P
Null pronoun	0.121	0.032	3.792	< .001**
Overt-pronoun-Ga	0.112	0.031	0.437	= .662
Overt-pronoun-Wa	0.113	0.029	2.171	= .030*

Note. Factors were coded with dummy coding. The subject antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

These comparisons indicate that null pronouns as well as repeated-name topic-*wa* were read significantly faster for subject antecedents than object antecedents. In contrast, reading times of repeated-name non-topic-*ga* were not significantly different when referring to subject and object antecedents. These outcomes are clearly reflected in the interactions between the anaphor form and antecedents. The OPP was only detected with overt-pronoun-*ga*, but not with overt-pronoun-*wa* because the overt-pronoun-*wa* showed similar antecedent preference to null pronouns.

Discussion

The results of Experiments 1 and 2 showed consistent outcomes in terms of the difference between the anaphors with *ga* and those with *wa*. The non-topic-*ga* anaphors showed the RNP and OPP, but the topic-*wa* anaphors showed neither of these effects. Regarding the non-topic anaphors with *ga* (both repeated names and overt pronouns), they were read significantly slower than null pronouns, but the reading-time differences between the non-topic-*ga* anaphors and null pronouns decreased significantly when the antecedent was the object, compared to when the antecedent was the subject. This interaction of anaphor (i.e., overt anaphors with *ga* vs. null anaphors) and antecedent salience (i.e., subject vs. object) shows that this is a discourse effect, and not merely a reflection of phonological, orthographic or any other baseline difference between the overt forms and null pronouns. If the outcomes were a reflection of the difference in phonological or orthographic lengths between null pronouns and overt anaphors, then the reading-time differences should have been uniform regardless of the grammatical role of the antecedent (i.e., subject vs. object), and no interaction between anaphor form and antecedent type would have been detected. The results of the non-topic-*ga* anaphors and null anaphors are attributed to their different sensitivities to the grammatical saliences of antecedents (i.e., subject antecedents vs. object antecedents). Specifically, while null pronouns are strongly sensitive to the different grammatical statuses of antecedents (as reflected in their reading-time difference depending on the antecedent's salience), non-topic-*ga* anaphors are significantly less sensitive. Almor's (1999) ILH is applicable to these results; for the salient antecedents, the non-topic repeated-name-*ga* and non-topic overt-pronoun-*ga* were processed slower than null pronouns because they were too

informative to identify the antecedents, but the informativity was to some extent justified when antecedents were non-salient because the rich information contributed to the identification of the antecedents. These results are similar to Spanish (Gelormini-Lezama & Almor, 2011), Brazilian-Portuguese and Italian (Almor, de Carvalho Maia, Cunha Lima, Gelormini-Lezama & Vernice, 2013), which showed both the RNP and OPP. One difference between Japanese and the above Romance languages is that processing a sentence with overt pronouns in the Romance languages is penalized by rich information carried by the overt pronouns (e.g., gender, number, and person) plus the information carried by verbal morphology. On the other hand, Japanese overt pronouns are too informative by themselves. As discussed in an earlier section, Japanese overt pronouns are different from the overt pronouns in English or Spanish in that Japanese overt pronouns connote information such as speaker's age and social statuses as well as the personal relationship and the psychological distance between the speaker and the addressee. These elements were likely redundant in order to identify salient (default) antecedents, and thus delayed their processing.

In contrast, the overt topic anaphors with *wa* and null pronouns did not show an interaction with the antecedents' saliences. These results are distinct from the non-topic anaphors with *ga*, and thus our results show clear evidence of the effect from anaphors' topic-hood. In other words, the topic anaphors with *wa* and the non-topic anaphors with *ga* are differently processed. Since these different outcomes were detected in spite of the same anaphor forms (i.e., repeated-name-*ga* vs. repeated-name-*wa*, and overt-pronoun-*ga* vs. overt-pronoun-*wa*) that carry the same amount of semantic information, Almor's (1999) argument does not explain this outcome. Instead, Gordon and Hendrick's (1998)

DPT may apply to this outcome. The DPT explains that an English overt pronoun is immediately interpreted as an anaphor and leads readers to search for an antecedent in decreasing order of antecedent prominence (i.e., salience), starting with the most prominent antecedent to the least prominent antecedent. In Japanese, as explained in an earlier section, a topic-NP with *wa* typically refers to information that is already acknowledged, while the NP with *ga* is a non-topic that generally introduces unacknowledged information in a discourse. In other words, the topic-postposition *wa* functions to signal that the NP-*wa* is an anaphor so that readers immediately search for its antecedent. Therefore, similar to null pronouns, the topic-*wa* anaphors allowed readers to find salient antecedents (i.e., subject) faster than non-salient antecedents (i.e., object), resulting in no RNP or OPP detected.

One thing to note is that, although both topic-*wa* anaphors did not exhibit the RNP or the OPP, there was a difference between the repeated-name topic-*wa* and overt-pronoun topic-*wa*. While repeated-name topic-*wa* showed only a marginally significant antecedent preference (Table 6.6), overt-pronoun topic-*wa* significantly preferred the subject antecedents to the object antecedents (Table 6.10), just like null pronouns did. These outcomes might reflect that, in addition to *wa*-marking, the overt pronoun form contributed to allowing readers to immediately interpret them as anaphors. Although Japanese overt pronouns are different from the overt pronouns in English or Spanish, they should be anaphorically interpreted as long as they are pronouns, which typically refer to a person that the speaker and addressee both know, regardless of whether they are affixed by *ga* or *wa*. In addition, this overt pronouns' similarity to null pronouns (in terms of their antecedent preferences) might reflect that, as Obana (2003b) observed, younger

Japanese speakers are beginning to use overt pronouns as mere pronouns without socio-cultural constraints, and our participants' ages were mostly under 20 ($M = 19.833$ in Experiment 1; $M = 19.974$ in Experiment 2).

As a possible limitation, I need to mention that the results of Experiments 1 and 2 are different from Shoji, Dubinsky and Almor (2015, 2016), which elicited the RNP and OPP with both the non-topic-*ga* and topic-*wa* anaphors; their results are distinct from the present study that found the RNP and OPP with the non-topic-*ga* anaphors but not with the topic-*wa* anaphors. A possible explanation could be attributed to the participants' exposure to the second language. In Shoji, Dubinsky and Almor's study, the participants were U.S. residents, and their lengths of residence in the U.S. ranged from one month to 22 years ($M = 5.84$), while the present study collected data from participants who had never continuously stayed outside Japan more than one month. The participants' language instinct in the former study might have been affected by English. They could have been less sensitive in differentiating the topic-subject and non-topic-subject because English does not overtly mark the topic-hood. In this situation, where the participants might have looked at only the subject-hood of anaphors, the RNP and OPP could have been easier to elicit even with the topic anaphor with *wa*, relative to the situation where all the participants are purer monolingual Japanese speakers (as in the present study).

Chapter 7. Experiments 3 and 4

Predictions

Experiments 3 and 4 investigated the following issues regarding the topic-hood of antecedents:

- (i) RNP: If the topic is more salient than the non-topic, participants may read sentences with null pronouns faster than sentences with repeated-name anaphors when the antecedents are the topic. However, the reading-time difference might significantly decrease when the antecedents are the non-topic.
- (ii) OPP: If the topic is more salient than the non-topic, participants may read sentences with null pronouns faster than sentences with overt-pronoun anaphors when the antecedents are the topic. However, the reading-time difference might significantly decrease when the antecedents are the non-topic.
- (iii) Topic: Participants may process the sentences with overt anaphors differently depending on whether they are the topic (i.e., repeated-name topic-*wa* and overt-pronoun topic-*wa*) or the non-topic (i.e., repeated-name non-topic-*ga* and overt-pronoun non-topic-*ga*).

If the topic-hood of the antecedents enhances the accessibility, sentences with null pronouns would be read faster than sentences with repeated-name/overt-pronoun anaphors when they refer to topic antecedents. However, the reading-time difference would be significantly decreased when they refer to non-topic (less salient) antecedents.

In that case, we would observe RNP and OPP with the topic antecedent and the non-topic antecedent.

Participants

Forty-one undergraduate students from the University of Shizuoka in Japan participated in Experiment 3. They were all native Japanese speakers and between 18 and 23 years of age ($M = 18.809$). Seventeen undergraduate students from the University of Shizuoka and twenty-six undergraduate students from Mie University (i.e., forty-two participants in total) participated in Experiment 4. They were all native Japanese speakers and between 18 and 21 years of age ($M = 18.524$). Each session lasted approximately 20 minutes.

Materials

Just like in Experiment 1, in Experiment 3's examination of the RNP, anaphors were repeated names with *ga* (i.e., repeated-name non-topic), repeated names with *wa* (i.e., repeated-name topic), and null pronouns. Just like in Experiment 2, in Experiment 4's examination of the OPP, anaphors were overt pronouns with *ga*, overt pronouns with *wa*, and null pronouns. In each experiment, two-sentence discourses were given, in which three types of anaphors refer to two different types of antecedents (topic vs. non-topic). Accordingly, there were six conditions in total in each experiment. Six items are provided for each condition, which creates 36 experimental items in each experiment. These are mixed among 48 distractors. Sample items are shown below.

Table 7.1

Conditions and Example Items for Experiment 3 (RNP)

Conditions:	Antecedent	Anaphor
	Subject-topic	Repeated name <i>-ga</i>
	Non-topic Subject	Repeated name <i>-wa</i>
		Null pronoun
<hr/>		
<u>Example items:</u>		
<hr/>		
First sentence with Antecedent		
(i) Topic Subject		
e.g.	<u>Taku-wa</u>	Kazuko-o arubaito tosite yatotta.
	Taku-TOP	Kazuko-ACC part-timer as hired
	‘Taku hired Kazuko as a part-timer.’	
(ii) Non-topic Subject		
e.g.	<u>Taku-ga</u>	Kazuko-o arubaito tosite yatotta.
	Taku-NOM	Kazuko-ACC part-timer as hired
	‘Taku hired Kazuko as a part-timer.’	
<hr/>		
Second sentence with Anaphor		
(i) Repeated name <i>-ga</i>		
e.g.	Taku-ga mainiti	Kazuko-to issmoni hataraita.
	Taku-NOM everyday	Kazuko-with together worked
	‘Taku worked with Kazuko everyday.’	
(ii) Repeated name <i>-wa</i>		
e.g.	Taku-wa mainiti	Kazuko-to issmoni hataraita.
	Taku-TOP everyday	Kazuko-with together worked
	‘Taku worked with Kazuko everyday.’	
(iii) Null pronoun		
e.g.	Ø mainiti	Kazuko-to issmoni hataraita.
	everyday	Kazuko-with together worked
	‘(he) worked with Kazuko everyday.’	

Table 7.2

Conditions and Example Items for Experiment 4 (OPP)

Conditions:	Antecedent	Anaphor
	Subject-topic	Overt pronoun <i>-ga</i>
	Non-topic Subject	Overt pronoun <i>-wa</i>
		Null pronoun

Example items:

First sentence with Antecedent

(i) Topic Subject

e.g. Taku-wa Kazuko-o arubaito tosite yatotta.
 Taku-TOP Kazuko-ACC part-timer as hired
 ‘Taku hired Kazuko as a part-timer.’

(ii) Non-topic Subject

e.g. Taku-ga Kazuko-o arubaito tosite yatotta.
 Taku-NOM Kazuko-ACC part-timer as hired
 ‘Taku hired Kazuko as a part-timer.’

Second sentence with Anaphor

(i) Overt pronoun *-ga*

e.g. Kare-ga mainiti Kazuko-to isshoni hataraita.
 he-NOM everyday Kazuko-with together worked
 ‘He worked with Kazuko everyday.’

(ii) Overt pronoun *-wa*

e.g. Kare-wa mainiti Kazuko-to isshoni hataraita.
 he-TOP everyday Kazuko-with together worked
 ‘He worked with Kazuko everyday.’

(iii) Null pronoun

e.g. Ø mainiti Kazuko-to isshoni hataraita.
 everyday Kazuko-with together worked
 ‘(he) worked with Kazuko everyday.’

All other aspects of items and procedure were the same as in Experiments 1 and 2.

Results: Experiment 3

The data from one participant whose accuracy was lower than 80% was removed prior to analysis. For the data from the remaining 40 participants, just like in Experiments 1 and 2, reading times from trials to which the participants incorrectly answered the comprehension question were removed, affecting 8.67% of the data. Also, extremely short or long reading times, which are below 300 ms or greater than 6000 ms, were removed, affecting 1.63%. The reading times of each type of anaphor are shown in Figure 7.1 below. Error bars in the figure represent the standard error of the mean.

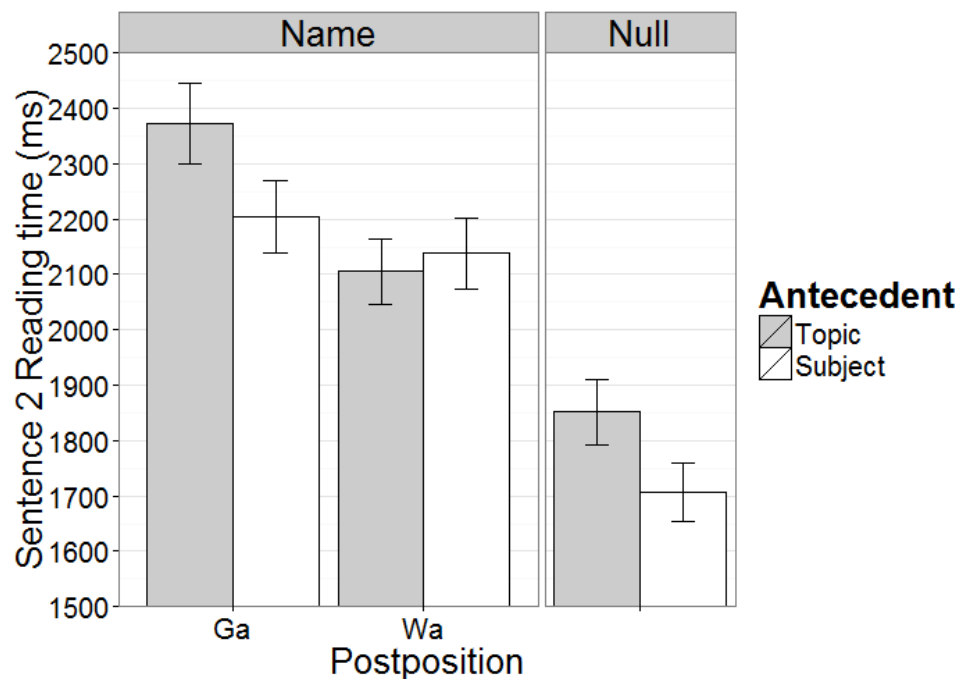


Figure 7.1. Reading times of the second sentences with anaphors in Experiment 3

The analysis of the data detected no significant interaction between antecedent (topic vs. non-topic) and anaphor (repeated-name-*ga*, repeated-name-*wa*, and null pronoun) ($\chi^2(2) = 3.93, p = .140$). Table 7.3 shows the coefficients of the fixed terms in the final model. The results were greatly different from our prediction in that null pronouns were read significantly slower when antecedents were the topic than when they were the non-topic subject. This outcome implies that the non-topic-subject antecedents are more salient than the topic-subject antecedents are. The coefficients of the interaction terms in this analysis were not significant with the repeated-name non-topic-subject-*ga*. The repeated-name topic-subject-*wa* showed a marginally significant interaction, which implies a weak RNP.

Table 7.3

Omnibus RNP Analyses for Experiment 3

	β	SE	T	P
Intercept	7.429	0.041	180.700	< .001**
Antecedent: Subject (Non-topic)	-0.068	0.034	-1.981	= .048*
Anaphor: Repeated-name-Ga	0.251	0.035	7.241	< .001**
Anaphor: Repeated-name-Wa	0.140	0.035	4.035	< .001**
Antecedent: Subject \times Anaphor: Repeated-name-Ga	0.002	0.049	0.035	= .972
Antecedent: Subject \times Anaphor: Repeated-name-Wa	0.085	0.049	1.740	= .082 .

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors,

respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

In addition, a series of analyses were conducted separately for each full reference form. The analyses contrasted repeated-name-*ga* vs. null pronouns and repeated-name-*wa* vs. null pronouns. The results are summarized in Tables 7.4 and 7.5 below. These analyses detected no interaction in repeated-name non-topic-*ga*, i.e., $\chi^2(1) = 0.004$, $p = .95$, but the interaction was marginally significant in repeated-name topic-*wa*, i.e., $\chi^2(1) = 3.09$, $p = .08$.

Table 7.4

Separate RNP Analysis for Experiment 3: Null Pronoun vs. Repeated-name Non-topic-ga

	β	SE	T	P
Intercept	7.429	0.041	183.015	< .001**
Antecedent: Subject (Non-topic)	-0.068	0.035	-1.969	= .049*
Anaphor: Repeated-name-Ga	0.251	0.035	7.197	< .001**
Antecedent: Subject \times Anaphor: Repeated-name-Ga	0.003	0.049	0.063	= .950

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent

differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 7.5

Separate RNP Analysis for Experiment 3: Null Pronoun vs. Repeated-name Topic-wa

	β	SE	T	P
Intercept	7.428	0.040	183.989	< .001**
Antecedent: Subject (Non-topic)	-0.066	0.034	-1.942	= .052 .
Anaphor: Repeated-name-Wa	0.141	0.034	4.111	< .001**
Antecedent: Subject \times Anaphor: Repeated-name-Wa	0.085	0.048	1.758	= .079 .

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The marginal interaction in repeated-name topic-wa indicated that the reading-time differences between repeated-name topic-wa and the null pronoun in the topic-antecedent conditions were smaller than in the non-topic-antecedent conditions, suggesting a marginal RNP. On the other hand, the reading-time difference between repeated-name non-topic-ga and null pronoun in the topic-antecedent condition was not different from

their reading-time difference in the non-topic-antecedent condition. Thus, there was no interaction between the anaphors and antecedents.

Like Experiments 1 and 2, a series of simple analyses comparing the reading time of each anaphor form between the topic and non-topic antecedent conditions was also conducted.

Table 7.6

Simple Topic vs. Non-topic Antecedent Effects for Each Anaphor Form in Experiment 3

	β	SE	T	P
Null pronoun	-0.068	0.034	-1.981	= .048*
Repeated-name-Ga	-0.066	0.035	-1.906	= .059 .
Repeated-name-Wa	0.017	0.035	0.481	= .630

Note. Factors were coded with dummy coding. The topic antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The comparisons indicate that null pronouns were read faster for the non-topic antecedent than the topic antecedent at the marginally significant level. On the other hand, repeated-name topic-*wa* did not show such an antecedent preference. As a result, they elicited the

marginal RNP. In addition, repeated-name non-topic-*ga* showed the same trend as null pronouns at a marginal significance, and thus these two types of anaphors did not show a RNP.

Results: Experiment 4

The data from one participant whose accuracy was lower than 80% was removed prior to analysis. For the remaining 41 participants' data, reading times from trials to which the participants incorrectly answered the comprehension questions were removed, which affected 7.14% of the data. In addition, extreme reading times below 300 ms or greater than 6000 ms were removed, affecting 1.28%. The reading times of each of the types of anaphors are shown in Figure 7.2 below. Error bars in the figure represent the standard error of the mean.

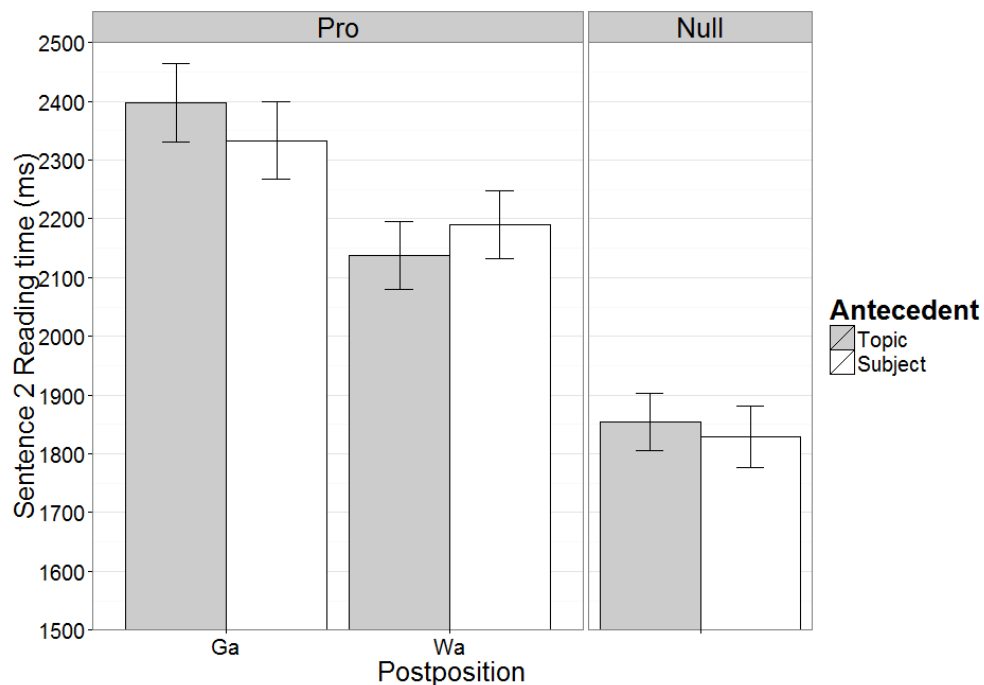


Figure 7.2. Reading times of the second sentences with anaphors in Experiment 4

Unlike Experiment 3, the analysis did not find any significant interaction between antecedent (topic vs. non-topic) and anaphor (overt-pronoun-*ga*, overt-pronoun-*wa*, and null pronoun) ($\chi^2(2) = 0.82, p = .662$). Table 7.7 shows the coefficients of the fixed terms in the final model. The coefficients of the interaction terms in this analysis were not significant in all the cases, suggesting no OPP.

Table 7.7

Omnibus OPP Analyses for Experiment 4

	β	SE	T	P
Intercept	7.441	0.042	175.642	< .001**
Antecedent: Subject (Non-topic)	-0.020	0.032	-0.631	= .528
Anaphor: Overt-pronoun-Ga	0.255	0.032	7.838	< .001**
Anaphor: Overt-pronoun-Wa	0.153	0.032	4.705	< .001**
Antecedent: Subject \times Anaphor: Overt-pronoun-Ga	0.007	0.046	-0.161	= .872
Antecedent: Subject \times Anaphor: Overt-pronoun-Wa	0.022	0.046	0.698	= .485

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Also, a series of analyses were separately conducted for each reference form. The analyses contrasted overt-pronoun-*ga* vs. null pronouns and overt-pronoun-*wa* vs. null pronouns. The results are summarized in Tables 7.8 and 7.9 below. Consistent with the omnibus analysis above, these analyses revealed that the interaction term was not necessary in both overt-pronoun non-topic-*ga*, i.e., $\chi^2(1) = 0.02$, $p = .88$, and overt-pronoun topic-*wa*, i.e., $\chi^2(1) = 0.47$, $p = .50$.

Table 7.8

Separate OPP Analysis for Experiment 4: Null Pronoun vs. Overt-pronoun Non-topic-ga

	β	SE	T	P
Intercept	7.440	0.043	173.603	< .001**
Antecedent: Subject (Non-topic)	-0.020	0.033	-0.611	= .541
Anaphor: Overt-pronoun-Ga	0.254	0.033	7.666	< .001**
Antecedent: Subject \times Anaphor: Overt-pronoun-Ga	-0.007	0.047	-0.146	= .884

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 7.9

Separate OPP Analysis for Experiment 4: Null Pronoun vs. Overt-pronoun Topic-wa

	β	SE	T	P
Intercept	7.441	0.043	174.831	< .001**
Antecedent: Subject (Non-topic)	-0.019	0.032	-0.618	= .537
Anaphor: Overt-pronoun-Wa	0.154	0.032	4.857	< .001**
Antecedent: Subject \times Anaphor: Overt-pronoun-Wa	0.031	0.045	0.682	= .495

Note. Factors were coded with dummy coding. The null pronoun and topic antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Moreover, as I conducted in previous experiments, simple comparisons contrasting the reading times of each anaphor form (null pronoun, overt-pronoun-*ga*, and overt-pronoun-*wa*) between the topic and non-topic antecedent conditions were conducted. The results are shown below.

Table 7.10

Simple Topic vs. Non-topic Antecedent Effects for Each Anaphor Form in Experiment 4

	β	SE	T	P
Null pronoun	-0.020	0.032	-0.631	= .528
Overt-pronoun-Ga	-0.028	0.033	-0.851	= .395
Overt-pronoun-Wa	0.012	0.033	0.358	= .720

Note. Factors were coded with dummy coding. The topic antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The coefficients corresponding to these simple comparisons indicate that none of the anaphors were read differently when paired with the topic or non-topic antecedents. Indeed, Experiment 4 did not elicit the OPP with any anaphors or antecedents. Also, a contradiction that we found is that null pronouns’ different reading times, depending on antecedent types detected in Experiment 3, were not found in Experiment 4.

Discussion

Unlike Experiments 1 and 2, whose antecedents were grammatically different (i.e., subject vs. object), Experiments 3 and 4 used antecedents that were different in terms of information structure (i.e., topic vs. non-topic). Experiment 3 elicited a marginal RNP

with the repeated-name topic-*wa* anaphors and null pronouns, but not with the repeated-name non-topic-*ga* anaphors. The difference between repeated-name topic-*wa* and repeated-name non-topic-*ga* indicates that *wa* and *ga* differently affected the processing of anaphors in the same form (i.e., repeated name). On the other hand, Experiment 4 did not elicit the OPP regardless of the postpositions, *wa* or *ga*.

Neither Experiment 3 nor 4 supported the prediction from Crawley (1986) and Colonna, Schimke and Hemforth (2012) that indicated that the topic-hood of the antecedents enhanced their saliences and accessibilities (in English and German, respectively, but not in French). On the contrary, Experiment 3 showed that null pronouns (and non-topic repeated-name anaphors with *ga*) were read faster when they referred to the non-topic-*ga* antecedent than when referring to the topic-*wa* antecedent (Table 7.6). This outcome implies that the non-topic-*ga* antecedent is more salient than the topic-*wa* antecedent. The null pronouns' preference for the non-topic subject-*ga* antecedent is in fact consistent with the corpus studies by Clancy and Downing (1987), Hinds (1984) and Nakahama (2003). They found that the most frequent combination of antecedents and anaphors in Japanese was the non-topic subject-*ga* antecedent referred to by null pronouns. In their studies, there were much fewer instances of the topic subject-*wa* antecedent being referred to with null pronouns.

Unlike null pronouns and repeated-name non-topic-*ga*, repeated-name topic-*wa* anaphors were read indifferently for either antecedent (Table 7.6), or numerically faster when referring to the topic-*wa* antecedents than when referring to the non-topic-*ga* antecedents. The different antecedent preferences of null pronouns and repeated-name topic-*wa* contributed to the elicitation of the marginal RNP. This marginally significant

interaction could be straightforwardly explained by the ILH. Null pronouns carry no semantic information, and thus they were read faster for more salient antecedents (i.e., non-topic antecedents) than for less salient antecedents (i.e., topic antecedents). On the other hand, repeated-name topic-*wa* anaphors were too informative for simply identifying already salient antecedents. Therefore, their reading was slowed, and their reading times for salient antecedents became closer to their reading times for non-salient antecedents. However, it should be noted that these results contradict with the results from Experiment 1. In Experiment 1, repeated-name-*wa* and null pronouns did not show an interaction with antecedents while they elicited a marginal interaction in Experiment 3. The different outcomes should be attributed to the difference of antecedents (subject vs. object in Experiment 1, and topic-subject vs. non-topic-subject in Experiment 3). Experiment 1 showed that null pronouns significantly preferred subject antecedents to object antecedents (Table 6.6, $p < .001$), which indicates that subject antecedents are significantly more salient than object antecedents. Although Experiment 3 found that null pronouns preferred non-topic-subject antecedents to topic-subject antecedents, this effect appeared to be weaker (Table 7.6, $p = .048$). This suggests that the difference of salience between topic and non-topic antecedents is weaker than that between subject and object antecedents. The difference of salience between topic and non-topic antecedents might have been too subtle for repeated-name-*wa* to detect, while null pronouns showed the effect albeit weakly in Experiment 3. Therefore, repeated-name topic anaphor with *wa* were read at similar speeds for topic antecedents and for non-topic antecedents, resulting in the marginally significant antecedent-anaphor interaction when anaphors were repeated-name-subject-*wa* and null pronouns.

Regarding the repeated-name non-topic-*ga* anaphors, similar to null pronouns, they were read faster for the non-topic-subject-*ga* antecedent than for the topic-subject-*wa* antecedent at a marginally significant level (Table 7.6). The reason why repeated-name non-topic-*ga* anaphors preferred the non-topic-*ga* antecedents to the topic-*wa* antecedents might reflect *wa* and *ga*'s functions. As discussed in an earlier section, a topic-initial sentence (that starts with an NP with *wa*) introduces a discourse topic and indicates that the following discourse will develop in relation to the topic entity. When readers processed the second sentences, the anaphors with *ga* initially looked like new information or a subtopic (i.e., *push* in Walker's Cache Model), which were the least likely to correspond with the discourse topic (i.e., antecedent-*wa* in the first sentence). Thus, the repeated-name non-topic-*ga* anaphor (that looked like a subtopic) would not have been interpreted as an anaphor for the topic-*wa* antecedents (that is the discourse topic), which resulted in the marginally significant delay of their reading times. In contrast, the repeated-name non-topic-*ga* anaphor did not particularly dis-prefer the non-topic-*ga* antecedents. In this case, an explicit discourse topic was absent in the entire discourse, and thus there was no matching between the discourse-topic antecedent and subtopic-like anaphor. Although they are not an appropriate anaphor-antecedent combination (as reflected in Experiment 1), processing of this combination (i.e., repeated-name-*ga* referring to the non-topic-*ga*) was not as inappropriate as the other combination (i.e., repeated-name-*ga* referring to the topic-*wa*).

Experiment 4 did not show an OPP; there was no interaction between the anaphor form and the antecedent's information-structure status. The failure to detect the interaction that was detected in Experiment 3 seems to be due to the different outcomes

of Experiment 3 and 4. While in Experiment 3 null pronouns showed a significant preference for the non-topic-*ga* antecedents (Table 7.6), this antecedent preference disappeared in Experiment 4 (Table 7.10). This discrepancy might be unexplainable because Experiments 3 and 4 used the exact same items, in terms of the null pronouns and antecedents.³³ Thus, the results of null pronouns' reading times should have been alike between Experiments 3 and 4. This inconsistency could be further tested with more participants. One possibility is that this null pronouns' preference detected in Experiment 3 might simply disappear with the testing of more participants, considering that the null pronouns' faster reading for subject antecedents than topic antecedents is not strongly significant. Also, this inconsistency could be attributed to the overt anaphors used in these experiments. While the repeated-name-*ga* in Experiment 3 activated readers' sensitivity to different antecedents at a marginally significant level (i.e., repeated-name-*ga* were read faster for subject antecedents than topic antecedents), the activation of their sensitivity to antecedents remained when they read null pronouns in Experiment 3. On the other hand, overt pronouns (both with *ga* and *wa*) in Experiment 4 did not show a reading-time difference between topic antecedents and non-topic antecedents. As a result, null pronouns were read faster for non-topic-*ga* antecedents than topic-*wa* antecedents in Experiment 3 but not in Experiment 4.

Another finding in Experiment 4 is that the overt-pronoun non-topic-*ga* anaphors (i.e., *kare/kanojo-ga*) were read indifferently between the topic-*wa* antecedents and the

³³ The only difference between Experiments 3 and 4 was that repeated names were used in Experiment 3, whereas overt pronouns were used in Experiment 4. Thus, the discourse items including null pronoun anaphors were exactly the same in these two experiments.

non-topic-*ga* antecedents (Table 7.10). This is unlike the repeated-name non-topic-*ga* anaphors in Experiment 3, which showed a marginally significant preference for the non-topic-*ga* antecedents (Table 7.6). Because they were marked by the same postposition *ga*, their difference in antecedent preferences should be attributed to the different anaphor forms, repeated name versus overt pronoun. In a previous paragraph, I explained that repeated-name non-topic-*ga* anaphors dis-preferred the topic-*wa* antecedents because the anaphors were initially interpreted as new information and as a subtopic. The fact that the overt-pronoun non-topic-*ga* did not show this dis-preference might indicate that they were interpreted, not as new information, but as an anaphor (i.e., old information), although they were marked by *ga*. As I explained for Experiments 1 and 2, a pronoun (either overt pronoun or null pronoun) should be interpreted as an anaphor by default (although Japanese pronouns are more than mere pronouns). Possibly the function of *ga* that typically introduces new information was neutralized by the overt pronoun form, resulting in the disappearance of the repeated-name-*ga* anaphors' dis-preference for the topic-*wa* antecedents.

Chapter 8. Experiments 5 and 6

Predictions

Experiments 5 and 6 investigate the following issues regarding empathy:

- (i) RNP: If the empathy locus is more salient than the non-empathy locus, participants may read sentences with null pronouns faster than sentences with repeated-name anaphors when the antecedents are the empathy locus. However, the reading-time difference might significantly decrease when the antecedents are the non-empathy locus.
- (ii) OPP: If the empathy locus is more salient than the non-empathy locus, participants may read sentences with null pronouns faster than sentences with overt pronouns when the antecedents are empathy locus. However, the reading-time difference might significantly decrease when the antecedents are the non-empathy locus.
- (iii) Topic: Participants may process sentences with the overt anaphors differently depending on whether they are the topic (i.e., repeated-name topic-*wa* and overt-pronoun topic-*wa*) or the non-topic (i.e., repeated-name non-topic-*ga* and overt-pronoun non-topic-*ga*).

If empathy enhances the accessibility of antecedents, sentences with null pronouns would be read faster than repeated-name/overt-pronoun anaphors when they refer to empathy-locus antecedents. However, the reading-time difference might be significantly neutralized when they refer to non-empathy-locus (less salient) antecedents. In that case, we would observe RNP and OPP with empathy-locus antecedents and non-empathy-locus antecedents. In addition, the anti-logophoricity of overt pronouns would lead to the

slower reading of overt-pronoun sentences preceded by empathy-locus antecedents compared to the reading of overt-pronoun sentences preceded by non-empathy-locus antecedents.

Participants

Forty-one undergraduate students from Nagoya University in Japan participated in Experiment 5. They were all native Japanese speakers and between 18 and 23 years of age ($M = 18.926$). Forty-two undergraduate students from Mie University participated in Experiment 6. They were all native Japanese speakers and between 18 and 24 years of age ($M = 19.523$). Each session lasted approximately 20 minutes.

Materials

Just like earlier experiments, two-sentence discourses were given, in which three types of anaphors referred to two antecedents (empathy locus vs. non-empathy locus). For Experiments 5 and 6, the empathy locus was expressed by the verbs *ageru/kureru* ‘give’ and *-te ageru/-te kureru* ‘do a favor of –ing’ as well as *motte-iku/motte-kuru* ‘bring’ and *katte-iku/katte-kuru* ‘buy’. These words had to appear in more than one experimental item. In order to minimize repetitions of the words, I used the *masu*-form (e.g., *agemasu/kuremasu*) and plain form (e.g., *ageru/kureru*) across verbs; the *masu*-form is used in more formal situations than the plain form. Moreover, I assigned different levels of politeness for *ageru*, which also created different forms of the verbs. Namely, I used *yaru* ‘give’ and *-te yaru* ‘do a favor of –ing’ along with *ageru* ‘give’ and *-te ageru* ‘do a favor of –ing’; *yaru* ‘give’ and *-te yaru* ‘do a favor of –ing’ are impolite versions of *ageru* ‘give’ and *-te ageru* ‘do a favor of –ing’. Overall, the variations of the verbs in the experimental items are summarized in the table below.

Table 8.1

Verbs in Items of Experiment 5 and 6³⁴

	Empathy locus = do-er (e.g., giver)	Empathy locus = do-ee (e.g., receiver)	
‘give’	ageru	kureru	(plain)
	agemasu	kuremasu	(formal)
	yaru		(plain, impolite)
	yarimasu		(formal, impolite)
‘do a favor of –ing’	-te ageru	-te kureru	(plain)
	-te agemasu	-te kuremasu	(formal)
	-te yaru		(plain, impolite)
	-te yarimasu		(formal, impolite)

³⁴ In Japanese, there are polite versions of *ageru/kureru*, which are *sasiageru/kudasaru*. However, these words were not used in Experiments 5 and 6 because overt pronouns are not used when referring to someone whose social status is higher than the speaker’s. That is, if a speaker says “Taro-ga... sasiageru/kudasaru”, that indicates that *Taro* is a social superior to the speaker, and thus the speaker should not use overt pronouns to refer to *Taro* (Hinds, 1975; Yashima, 2014), e.g., a discourse such as “Taro_i-ga... sasiageta. Kare_i...” is unacceptable.

‘bring’	motte iku	motte kuru	(plain)
	motte ikimasu	motte kimasu	(formal)
‘buy (for someone)’	katte iku	katte kuru	(plain)
	katte ikimasu	katte kimasu	(formal)

The number of items above varied: (i) discourses including ‘give’ in their first sentences were 8 items, (ii) discourses including ‘do a favor of –ing’ were 20 items, (iii) discourses including ‘bring’ were 4 items, and (iv) discourses including ‘buy’ were 4 items. Among the items above, the numbers of (iii) ‘bring’ and (iv) ‘buy’ were kept minim because these verbs had to appear in the exactly same form. The number of (i) ‘give’ was a little more than ‘bring’ and ‘buy’ because ‘give’ that indicates givers as the empathy locus has two variations, *ageru* and its impolite version *yaru*. The number of (ii) ‘do a favor of –ing’ was the largest because in this category *ageru/yaru* ‘give’ and *kureru* ‘give’ were attached to other verbs, i.e., *osiete-ageru* ‘do a favor of teaching’ or *okutte-ageru* ‘do a favor of sending’. By combining with different verbs, I expected that the effect from the repetition of *ageru/yaru* and *kureru* would be neutralized.

In addition, in order to distract participants from the repetition of the words across the experimental items, I prepared a larger number of distractor items than in Experiments 1, 2, 3 and 4; namely, 84 distractors. Six experimental items for each of the six conditions create 36 items, which are mixed among the 84 distractors (Total = 120 items).

Table 8.2

Conditions and Example Items for Experiment 5 (RNP)

Conditions:	Antecedent	Anaphor
	Empathy-locus Subject	Repeated name -ga
	Non-empathy-locus Subject	Repeated name -wa
		Null pronoun

Example items:

First sentence with Antecedent

(i) Empathy-locus Subject

e.g. Taku-ga Kazuko-ni kuruma-o utte-ageta.
Taku-NOM Kazuko-DAT car-ACC sold
‘Taku sold his car for Kazuko.’ [Taku = Empathy locus]

(ii) Non-empathy-locus Subject

e.g. Taku-ga Kazuko-ni kuruma-o utte-kureta.
Taku-NOM Kazuko-DAT car-ACC sold
‘Taku sold his car for Kazuko.’ [Kazuko = Empathy locus]

Second sentence with Anaphor

(i) Repeated name -ga

e.g. Taku-ga Kazuko-kara daikin-o ikkatu-de uketotta.
Taku-NOM Kazuko-from money-ACC at once received
‘Taku received the money from Kazuko all at once.’

(ii) Repeated name -wa

e.g. Taku-wa Kazuko-kara daikin-o ikkatu-de uketotta.
Taku-TOP Kazuko-from money-ACC at once received
‘Taku received the money from Kazuko all at once.’

(iii) Null pronoun

e.g. Ø Kazuko-kara daikin-o ikkatu-de uketotta.
 Kazuko-from money-ACC at once received
‘(he) received the money from Kazuko all at once.’

Table 8.3

Conditions and Example Items for Experiment 6 (OPP)

Conditions:	Antecedent	Anaphor
	Empathy-locus Subject	Overt pronoun <i>-ga</i>
	Non-empathy-locus Subject	Overt pronoun <i>-wa</i>
		Null pronoun
<hr/>		
<u>Example items:</u>		
<hr/>		
First sentence with Antecedent		
(i) Empathy-locus Subject		
e.g.	<u>Taku-ga</u> Kazuko-ni kuruma-o utte-ageta.	
	Taku-NOM Kazuko-DAT car-ACC sold	
	‘Taku sold his car for Kazuko.’ [Taro = Empathy locus]	
(ii) Non-empathy-locus Subject		
e.g.	<u>Taku-ga</u> Kazuko-ni kuruma-o utte-kureta.	
	Takju-NOM Kazuko-DAT car-ACC sold	
	‘Taku sold his car for Kazuko.’ [Kazuko = Empathy locus]	
<hr/>		
Second sentence with Anaphor		
(i) Overt pronoun <i>-ga</i>		
e.g.	Kare-ga Kazuko-kara daikin-o ikkatu-de uketotta.	
	he-NOM Kazuko-from money-ACC at once received	
	‘He received the money from Kazuko all at once.’	
(ii) Overt pronoun <i>-wa</i>		
e.g.	Kare-wa Kazuko-kara daikin-o ikkatu-de uketotta.	
	he-TOP Kazuko-from money-ACC at once received	
	‘He received the money from Kazuko all at once.’	
(iii) Null pronoun		
e.g.	Ø Kazuko-kara daikin-o ikkatu-de uketotta.	
	Kazuko-from money-ACC at once received	
	‘(He) received the money from Kazuko all at once.’	

Just like Experiment 1 through 4, the number of letters and morae were kept approximately the same. For the sentences with null pronouns, the number of letters and morae ranged from 14 to 16 and 17 to 20, respectively. For the sentences with overt pronouns and repeated names, they ranged from 16 to 19 and from 21 to 24, respectively. A difference between Experiments 5 and 6 and Experiments 1 through 4 is that the personal entities as grammatical objects in the experimental discourses are indirect objects (with the dative-postposition *ni*) in Experiments 5 and 6, whereas in the other experiments they were direct objects (with the accusative-postposition *o*). In Experiments 5 and 6, with the verbs ‘give’, ‘do a favor of –ing’, ‘bring’ and ‘buy’, the direct objects had to be things to be given/brought/bought, and thus the recipient personal entities had to be indirect objects. All other aspects of items and procedure were the same as in Experiments 1 through 4.

Results: Experiment 5

The reading times of the second sentences that included anaphors were measured. Again, of the data from the 41 participants, reading times from trials to which the participant answered the comprehension question incorrectly were removed, affecting 9.01% of the data.³⁵ I removed extreme reading times below 300 ms or greater than 6000

³⁵ The incorrect responses for comprehension questions were relatively higher in Experiments 5 and 6 than the other experiments. This might reflect that the first sentences in Experiments 5 and 6, which included three arguments (i.e., subject, direct object, indirect object), might have been more complex and thus difficult to comprehend than those in the other experiments that included two arguments (subject, direct object).

ms, which affected 0.45% of the data. The reading times of anaphors are shown in Figure 8.1 below. Error bars in the figure represent the standard error of the mean.

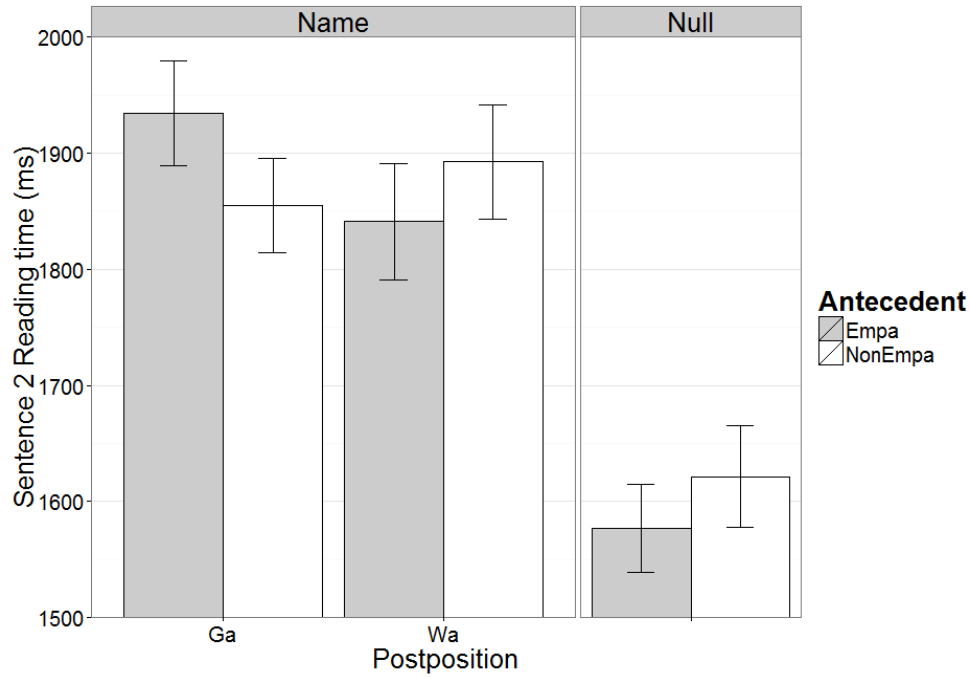


Figure 8.1. Reading times of the second sentences with anaphors in Experiment 5

There was no significant interaction between antecedent (empathy locus vs. non-empathy locus) and anaphor (repeated-name-*ga*, repeated-name-*wa*, and null pronoun) ($\chi^2(2) = 2.69, p = .261$) detected in the analysis. Table 8.4 shows the coefficients of the fixed terms in the final model. The coefficients of the interaction terms in this analysis proved to be not significant in all cases.

Table 8.4

Omnibus RNP Analyses for Experiment 5

	β	SE	T	P
Intercept	7.284	0.046	158.128	< .001**
Antecedent: Non-empathy-locus	0.022	0.027	0.816	= .415
Anaphor: Repeated-name-Ga	0.020	0.027	7.480	< .001**
Anaphor: Repeated-name-Wa	0.149	0.027	5.441	< .001**
Antecedent: Non-empathy \times				
Anaphor: Repeated-name-Ga	-0.053	0.038	-1.369	= .171
Antecedent: Non-empathy \times				
Anaphor: Repeated-name-Wa	0.004	0.039	0.103	= .918

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Like the other experiments, I also conducted the analyses that aimed to detect the RNP for each full reference form separately. The analyses compared repeated-name-*ga* vs. null pronouns and repeated-name-*wa* vs. null pronouns. The results are shown below.

These analyses found the interaction neither when anaphors included repeated-name non-topic-*ga*, i.e., $\chi^2(1) = 0.18$, $p = .18$, nor when anaphors included repeated-name topic-*wa*, i.e., $\chi^2(1) < .001$, $p = .98$.

Table 8.5

Separate RNP Analysis for Experiment 5: Null Pronoun vs. Repeated-name Non-topic-ga

	<i>B</i>	<i>SE</i>	<i>T</i>	<i>P</i>
Intercept	7.284	0.045	161.419	< .001**
Antecedent: Non-empathy-locus	0.022	0.027	0.836	= .403
Anaphor: Repeated-name-Ga	0.201	0.027	7.594	< .001**
Antecedent: Non-empathy ×				
Anaphor: Repeated-name-Ga	-0.050	0.038	-1.334	= .182

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 8.6

Separate RNP Analysis for Experiment 5: Null Pronoun vs. Repeated-name Topic-wa

	β	SE	T	P
Intercept	7.283	0.047	154.983	< .001**
Antecedent: Non-empathy-locus	0.024	0.028	0.869	= .385
Anaphor: Repeated-name-Wa	0.152	0.028	5.440	< .001**
Antecedent: Non-empathy \times Anaphor: Repeated-name-Wa	0.001	0.039	0.019	= .985

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

In addition, I compared the reading time of each anaphor form between the empathy-locus antecedent condition and non-empathy-locus antecedent condition.

Table 8.7

Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 5

	β	SE	T	P
Null pronoun	0.022	0.027	0.816	= .415
Repeated-name-Ga	-0.030	0.027	-1.120	= .263
Repeated-name-Wa	0.026	0.027	0.953	= .341

Note. Factors were coded with dummy coding. The empathy-locus antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The comparisons indicate that none of the anaphors were read differently when paired with the empathy-locus and non-empathy-locus antecedents. Although these results did not show any significance, our prediction could be supported by the trend that reading times of repeated-name-*ga* were larger with the empathy-locus antecedents than the non-empathy-locus antecedents, while null pronouns’ reading times were longer with the non-empathy-locus antecedents than those with the empathy-locus antecedents.

Results: Experiment 6

Of the data from the 42 participants, I removed the reading times from trials to which the participant incorrectly answered the comprehension question, which affected

10.58% of the data. Extreme reading times below 300 ms or greater than 6000 ms were removed, affecting 1.33% of the data. The reading times for each of the types of anaphors are shown in Figure 8.2 below. Error bars in the figure represent the standard error of the mean.

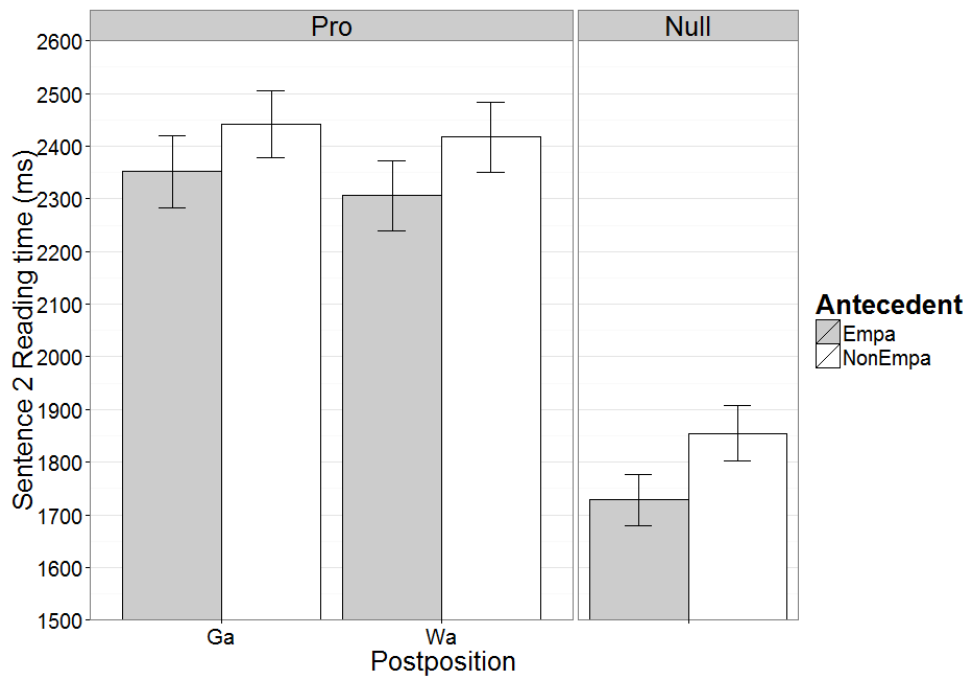


Figure 8.2. Reading times of the second sentences with anaphors in Experiment 6³⁶

³⁶ The raw reading times of Experiment 6 were slower than Experiment 5, which might reflect a difference between the participants in Experiments 5 (students of Nagoya University) and 6 (students of Mie University). Nagoya University is academically ranked higher than Mie University, (<http://daigaku.ehoh.net/hensachi/index.html>), so the students of Nagoya University could have processed information faster than the students of Mie University, similar to Experiments 1 and 2.

The analysis showed no significant interaction between antecedent (empathy locus vs. non-empathy locus) and anaphor (overt-pronoun-*ga*, overt-pronoun-*wa*, and null pronoun) ($\chi^2(2) = 0.34, p = .843$). Table 8.8 shows the coefficients of the fixed terms in the final model. The coefficients of the interaction terms in this analysis were not significant in all cases, suggesting no OPP.

Table 8.8

Omnibus OPP Analyses for Experiment 6

	β	SE	T	P
Intercept	7.360	0.049	151.200	< .001**
Antecedent: Non-empathy-locus	0.072	0.034	2.108	= .035*
Anaphor: Overt-pronoun-Ga	0.287	0.035	8.243	< .001**
Anaphor: Overt-pronoun-Wa	0.275	0.034	8.009	< .001**
Antecedent: Non-empathy ×				
Anaphor: Overt-pronoun-Ga	-0.021	0.049	-0.432	= .666
Antecedent: Non-empathy ×				
Anaphor: Overt-pronoun-Wa	-0.027	0.049	-0.555	= .579

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are

marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Just like the other experiments, I compared each full reference form separately. The analyses compared overt-pronoun-*ga* vs. null pronouns and overt-pronoun-*wa* vs. null pronouns. The results are summarized in Tables 8.9 and 8.10 below. These analyses found no interaction in either overt-pronoun non-topic-*ga*, i.e., $\chi^2(1) = 0.25$, $p = .62$, or in overt-pronoun topic-*wa*, i.e., $\chi^2(1) = 0.30$, $p = .59$.

Table 8.9

Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Non-topic-ga

	β	SE	T	P
Intercept	7.362	0.048	152.502	< .001**
Antecedent: Non-empathy-locus	0.074	0.033	2.229	= .026*
Anaphor: Overt-pronoun-Ga	0.288	0.034	8.476	< .001**
Antecedent: Non-empathy \times Anaphor: Overt-pronoun-Ga	-0.024	0.048	-0.502	= .616

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are

marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 8.10

Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Topic-wa

	β	SE	T	P
Intercept	7.360	0.047	156.106	< .001**
Antecedent: Non-empathy-locus	0.071	0.034	2.077	= .038*
Anaphor: Overt-pronoun-Wa	0.275	0.034	7.980	< .001**
Antecedent: Non-empathy \times Anaphor: Overt-pronoun-Wa	-0.027	0.049	-0.546	= .585

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Simple comparisons contrasting the reading times of each anaphor form between empathy-locus and non-empathy-locus antecedent conditions were conducted. The results are shown below.

Table 8.11

Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 6

	β	SE	T	P
Null pronoun	0.072	0.034	2.108	= .035*
Overt-pronoun-Ga	0.051	0.035	1.446	= .148
Overt-pronoun-Wa	0.045	0.034	1.323	= .186

Note. Factors were coded with dummy coding. The empathy-locus antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

These comparisons indicate that null pronouns were read significantly faster when referring to the empathy-locus rather than the non-empathy-locus antecedent. On the other hand, neither type of overt pronouns elicited significant reading-time differences. Nevertheless, we did not detect the OPP. Also, a contradiction is that, while Experiment 5 did not elicit significant reading-time differences with null pronouns when paired with the empathy-locus antecedents and non-empathy-locus antecedents, null pronouns in Experiment 6 did elicit a significant reading-time difference. This contradiction is similar to that in Experiments 3 and 4.

Results: Experiment 6 with *ageru* and *kureru*

Additional analyses were conducted for Experiments 5 and 6. I performed the same analyses from the previous section with the data reduced based on the type of verbs in the experimental items. In the first sentences of the discourses for Experiments 5 and 6, four different types of verbs were used: ‘give’ *ageru* (and *youtu*) vs. *kureru*, ‘do a favor of –ing’ *-te ageru* (and *-te yaru*) vs. *-te kureru*, ‘bring’ *motte-iku* vs. *motte-kuru*, and ‘buy’ *katte-iku* vs. *katte-kuru*. The verbs indicated whether the antecedent was the empathy locus or non-empathy locus. These verbs were divided into two groups; (i) the verbs that included *ageru* (and *youtu*)/*kureru* ‘give’ and ‘do a favor of –ing’ and (ii) the verbs that included *iku/kuru* ‘bring’ and ‘buy’. The additional analyses shown in this section were conducted with only (i) the verbs that included *ageru* (and *youtu*)/*kureru*. After removing (ii) the verbs that included *iku/kuru*, the data from 28 discourse items were left to be analyzed.³⁷

The verbs (ii) ‘bring’ *motte-iku/kuru* and ‘buy’ *katte-iku/kuru* were removed from the analyses because, although *iku/kuru* ‘come/go’ are widely acknowledged as empathy indicators (Kawakami, 1996; Kuno, 1978; Koga, 2014), no existing studies (to my knowledge) have argued that those verbs indicate empathy when they are used auxiliarily as *motte-iku/kuru* ‘bring’ or *katte-iku/kuru* ‘buy’. On the other hand, the verbs (ii) *ageru/kureru* ‘give’ and *-te ageru/kureru* ‘do a favor of –ing’ were analyzed because

³⁷ I did not conduct the additional analyses for (ii) the verbs that include *iku/kuru* (*motte-iku/kuru* ‘bring’ and *katte-iku/kuru* ‘buy’) because there were only 8 discourse items with these verbs (for 6 conditions), which may be too few to be statistically analyzed.

they are generally accepted as empathy-indicating devices, which is empirically supported by Koga's (2014) questionnaire-based research.

In the additional analyses with the selected data from the discourses that included *ageru/kureru*, Experiment 5 still showed no interaction between antecedents and anaphor forms in any case, indicating no RNP. However, in the analyses for Experiment 6, marginal interactions were detected when anaphors were overt pronouns (both with *ga* and *wa*), suggesting an OPP. The reading times of the overt-pronoun anaphors (in the second sentences) preceded by *ageru/kureru* (in the first sentences) are shown in Figure 8.3 below. Error bars in the figure represent the standard error of the mean.

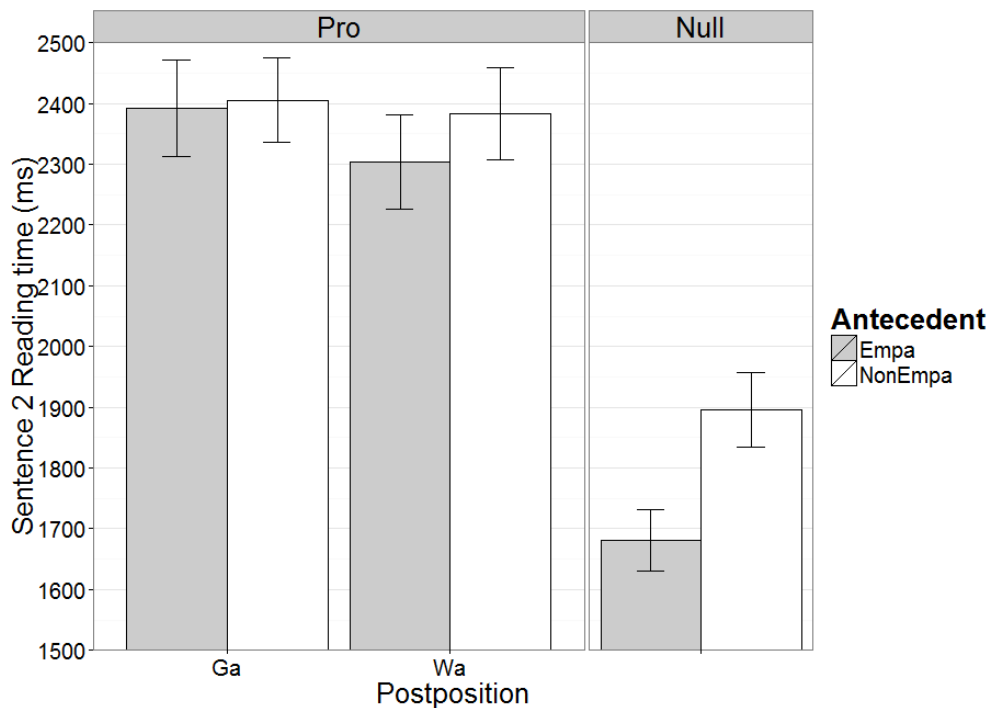


Figure 8.3. Reading times of the second sentences with anaphors in Experiment 6 with items *ageru/yaru* & *kureru*

The omnibus analysis of the selected data set found no interaction between antecedent (empathy locus and non-empathy locus) and anaphor (overt-pronoun-*ga*, overt-pronoun-*wa*, and null pronoun) ($\chi^2(2) = 4.09, p = .130$), but the coefficients of the fixed items in the final model (shown below) detected marginally significant interactions both with overt-pronoun-*ga* and overt-pronoun-*wa*.

Table 8.12

Omnibus OPP Analyses for Experiment 6 with Ageru/Kureru ‘give’ & ‘do a favor of –ing’

	β	SE	T	P
Intercept	7.334	0.051	144.849	< .001**
Antecedent: Non-empathy-locus	0.125	0.039	3.233	= .001**
Anaphor: Overt-pronoun-Ga	0.325	0.039	8.262	< .001**
Anaphor: Overt-pronoun-Wa	0.300	0.038	7.803	< .001**
Antecedent: Non-empathy ×				
Anaphor: Overt-pronoun-Ga	-0.099	0.056	-1.767	= .077 .
Antecedent: Non-empathy ×				
Anaphor: Overt-pronoun-Wa	-0.094	0.055	-1.724	= .085 .

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are

marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

One possible issue in these analyses is a small sample size due to the reduced items. Thus, I repeated the same analyses without items as a random factor. These repeated analyses did not change the results. Marginal interactions were detected with overt-pronoun-*ga* ($\beta = -0.102$, $SE = 0.059$, $t = -1.746$, $p = .086$) and with overt-pronoun-*wa* ($\beta = -0.099$, $SE = 0.057$, $t = -1.720$, $p = .086$).

In an effort to detect an OPP for each full reference separately, I compared the overt-pronoun-*ga* vs. null pronoun and overt-pronoun-*wa* vs. null pronoun. Marginally significant interactions between antecedent and anaphor were found with both overt-pronoun-*ga* ($\chi^2(1) = 3.141$, $p = .076$) and overt-pronoun-*wa* ($\chi^2(1) = 3.118$, $p = .077$).

Table 8.13

Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Non-topic-ga with Ageru/Kureru ‘give’ & ‘do a favor of –ing’

	<i>B</i>	<i>SE</i>	<i>T</i>	<i>P</i>
Intercept	7.336	0.050	146.624	< .001**
Antecedent: Non-empathy-locus	0.126	0.038	3.359	< .001**
Anaphor: Overt-pronoun-Ga	0.322	0.038	8.411	< .001**
Antecedent: Non-empathy ×				
Anaphor: Overt-pronoun-Ga	-0.096	0.054	-1.774	= .076 .

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

Table 8.14

Separate OPP Analysis for Experiment 6: Null Pronoun vs. Overt-pronoun Topic-wa with Ageru/Kureru ‘give’ & ‘do a favor of –ing’

	β	SE	T	P
Intercept	7.334	0.048	152.347	< .001**
Antecedent: Non-empathy-locus	0.126	0.039	3.246	= .001*
Anaphor: Overt-pronoun-Wa	0.301	0.039	7.778	< .001**
Antecedent: Non-empathy × Anaphor: Overt-pronoun-Wa	-0.097	0.055	-1.768	= .077 .

Note. Factors were coded with dummy coding. The null pronoun and empathy-locus antecedent were used as reference levels (value = 0) for the reference form and antecedent factors, respectively. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

As in the previous analysis, due to a concern regarding sample size, I repeated these analyses without items as a random factor, which resulted in similar outcomes. There still were marginally significant interactions for overt-pronoun-*ga* ($\beta = -0.100$, $SE = 0.057$, $t = -1.743$, $p = .082$) and for overt-pronoun-*wa* ($\beta = -0.100$, $SE = 0.057$, $t = -1.739$, $p = .082$).

In addition, I compared the reading time of each anaphor form between the empathy-locus and non-empathy-locus antecedent conditions, as shown below.

Table 8.15

Simple Empathy-locus vs. Non-empathy-locus Antecedent Effects for Each Anaphor Form in Experiment 6 with Ageru/Kureru ‘give’ & ‘do a favor of-ing’

	β	SE	T	P
Null pronoun	0.125	0.039	3.233	= .001**
Overt-pronoun-Ga	0.026	0.040	0.651	= .515
Overt-pronoun-Wa	0.031	0.039	0.790	= .430

Note. Factors were coded with dummy coding. The empathy-locus antecedent was always used as reference levels (value = 0) for antecedent factor. Reading times were log transformed and the coefficients thus represent differences between conditions in log scale. Effects at a $p \leq .05$ level are marked with a * and at a $p \leq .001$ level with a **. Effects with $p < .1$ are marked with a ‘.’.

The analyses for simple antecedent effects for each anaphor forms above show that, (when empathy locus was indicated by *ageru/kureru*-type of verbs) null pronouns were read significantly faster when antecedents were the empathy locus than when they were the non-empathy locus. On the other hand, there was no difference in reading times of the overt-pronoun-*ga/wa* between the two antecedent types. These different sensitivities for antecedents between null pronouns and overt-pronoun-*ga/wa* elicited the marginally significant OPP.

Discussion

Experiment 5 failed to elicit the RNP with all the analyses that I conducted, which might mean that repeated names are not significantly different from null pronouns in their sensitivities to an antecedents' empathy status (empathy locus vs. non-empathy locus). My prediction that repeated names would tend to prefer non-salient antecedents, i.e., non-empathy-locus antecedents, compared to null pronouns could be only supported by the trend that the repeated-name non-topic-*ga* showed. Notice that, in Experiment 5, sentences with the repeated-name-*ga* was numerically read faster when antecedents were the non-empathy locus than when they were the empathy locus, while null pronouns showed the opposite trend, although their different antecedent-preferences did not reach the necessary level to show a significant interaction between them and antecedents ($\beta = -0.053$, $SE = 0.038$, $t = -1.369$, $p = .171$). On the other hand, repeated-name topic-*wa* did not show any such trend, and they were read similarly to null pronouns, i.e., repeated-name topic-*wa* (similar to null pronouns) were read numerically faster for empathy-locus antecedents than for non-empathy-locus antecedents ($\beta = 0.004$, $SE = 0.039$, $t = 0.103$, $p = .918$). These results are somewhat similar to Experiment 1. The repeated names with *ga*

were processed differently from null pronouns (which we predicted to elicit the RNP) possibly because repeated names carry rich semantic information that is more than required to identify the salient (empathy-locus) antecedent, in accordance with the ILH. In contrast, the topic-hood given by *wa* neutralized the delay of reading by allowing readers to immediately interpret the repeated-name-*wa* as an anaphor, as the DPT argues.

Experiment 6 did not elicit the OPP when the entire dataset was analyzed. However, a marginal OPP was found when analyses were conducted only with the data from selected discourses including *ageru/kureru*-types of verbs (i.e., ‘give’ and ‘do a favor of –ing’). The results showed typical OPP-like outcomes (although the interactions were marginal). This supports the hypothesis that empathy-locus antecedents are more salient than non-empathy-locus antecedents, as Walker, Iida and Cote (1994) argue. Almor’s (1999) ILH can explain these results in a straightforward manner, which is supported by the Japanese overt pronouns’ rich semanticities. While null pronouns that carry no semantic information match the default antecedents (salient, empathy-locus antecedents), Japanese overt pronouns that contain rich semanticity are too informative to identify the default antecedents, which resulted in the marginal OPP.

Gordon and Hendrick’s (1998) DPT does not predict these effects; regardless of the postposition (*ga* vs. *wa*), overt pronouns showed the OPP, being read indifferently between with empathy-locus antecedent and non-empathy-locus antecedents. Recall that Experiments 1 and 2 revealed that topic anaphors with *wa* led readers to immediately interpret them as anaphors and were not subject to the RNP or OPP, as predicted by the DPT. The effect of *wa* disappeared in Experiment 6 (with selected verb items). The neutralization of *wa*’s function could be attributed to the anti-logophoricity of Japanese

overt pronouns. As discussed in Chapter 2, Japanese overt pronouns tend to avoid referring to a perspective-bearer, which may correspond with the empathy locus, according to Sells (1987). The anti-logophoricity and rich semanticity could additively contribute to the marginal OPP even with the topic anaphors with *wa*.

However, if it is true that Japanese overt pronouns carry anti-logophoricity, sentences with those should have been read faster for non-empathy-locus antecedents (non-perspective-bearer) than for empathy-locus antecedents (perspective-bearer). In reality, the effect of the anti-logophoricity was not strong enough to generate this outcome; overt pronouns were read indifferently between with empathy-locus antecedents and non-empathy-locus antecedents. This absence of the obvious anti-logophoricity in Experiment 6 could be because the discourses used in the experiment did not reliably elicit the anti-logophoricity. Compared to the present study, earlier studies related to anti-logophoricity more explicitly indicated empathy-loci by using the phrases such as “According to [empathy locus], he...” (Dubinsky & Hamilton, 1999). While our items attempted to indicate the empathy locus by using the verbs, *ageru/kureru* (and *iku/kuru*), these might have been too implicit to fully activate overt pronouns’ anti-logophoricity in readers’ processing. A future study could test the anti-logophoricity of Japanese overt pronouns more directly by using these phrases (i.e., “according to [empathy-locus entity]...” or “[empathy-locus entity] thinks that...”), which might elicit the OPP at a significant level. Moreover, as mentioned in the Materials section for Experiment 5 and 6, one limitation of Experiments 5 and 6 was that the same verbs had to be used several times in the experimental discourses because there are not many verbs that indicate empathy loci. There is a possibility that the participants became familiar

with the verbs while they were taking the experiments and became less reactive to the antecedents' empathy status. A solution for this possible problem could be conducting the experiment using fewer discourses which allows for decreased repetition of the same verbs, and in turn, an increase in the number of participants will be needed.

In Experiments 5 and 6, there is a similar problem to that of Experiments 3 and 4. Null pronouns' reading times showed different results between Experiments 5 and 6, although they used exactly the same items on the conditions with null pronouns. Specifically, null pronouns did not show a significant difference between empathy-locus and non-empathy-locus antecedents in Experiment 5 (Table 8.7), but a significant difference was detected in Experiment 6 (Table 8.11), favoring for the empathy-locus antecedents. This discrepancy could be attributed to the same reason for the inconsistency between Experiments 3 and 4: the overt anaphors used in Experiments 5 and 6 affected the reading of null pronouns. While overt pronouns (with *ageru/kureru* 'give') in Experiment 6 activated readers' sensitivity to the antecedent difference at a marginally significant level (i.e., sentences with overt pronouns were read faster for empathy-locus antecedents than for non-empathy-locus antecedents), their sensitivity to antecedents kept activated even when they read null pronouns in Experiment 6. On the other hand, any overt anaphors in Experiment 5 did not show a reading-time difference between two types of anaphors. Therefore, the readings times of null pronouns faster for empathy-locus antecedents than non-empathy-locus antecedents in Experiment 6 but not in Experiment 5. A supporting data for this explanation might lie in the data of Experiment 6. When I analyzed the data of Experiment 6 only with *ageru/kureru*-discourses, the null pronouns reading-time difference increased (Table 8.15, $p = .001$) compared to when I

analyzed the entire data including all the verbs in the first sentence (Table 8.11, $p = .035$).

This outcome indicates that, as the combination of overt pronouns and *ageru* and *kureru* ('give') made readers aware of the different salience of antecedents, the participants' awareness to different antecedents' salience remained for null pronouns as well.

Chapter 9. General Discussion

An original concern of this dissertation regarding the early studies of the RNP and OPP was that the languages that have been tested (e.g., English, Spanish) do not explicitly mark the topic-hood of anaphors. Since the topic-hood and the subject-hood generally overlap in those languages, it was questioned whether the non-topic subject anaphors and the topic subject anaphors might elicit different effects. Therefore, the current study investigated the RNP and OPP in Japanese, which morphologically marks the topic in order to separately examine the topic anaphors and the non-topic anaphors. The topic-postposition *wa* marked the overt topic anaphors, and the nominative-postposition *ga* marked the non-topic anaphors. These overt anaphors were used, as well as null pronouns, to find out preferences for and sensitivities to different types of antecedents. Since overt anaphors were either repeated names or overt pronouns, five types of anaphors were utilized in total: (i) repeated-name non-topic-subject-*ga*, (ii) repeated-name topic-subject-*wa*, (iii) overt-pronoun non-topic-subject-*ga*, (iv) overt-pronoun topic-subject-*wa*, and (v) null pronoun (that is also the topic).

Another concern of this study was about the salience of the antecedent. While many existing studies confirm that the grammatical subject is more salient than the grammatical object, a number of studies propose that the topic antecedents might be more salient than the non-topic antecedents and that the empathy-locus antecedents could be more salient than the non-empathy-locus antecedents. Therefore, while earlier literatures controlled antecedent saliences only by their grammatical statuses and word order, this

study attempted to manipulate the antecedents' saliences by information-structural status (topic vs. non-topic, Experiments 3 and 4) and a non-linguistic notion, namely the empathy status (empathy locus vs. non-empathy locus, Experiments 5 and 6), in addition to grammatical status (subject vs. object, Experiments 1 and 2). These possibly different saliences of antecedents were expected to elicit interactions with different anaphor types.

Let us review two earlier accounts for the RNP/OPP, namely the DPT and the ILH, which are used as the frameworks for explaining the results of the current study. Gordon and Hendrick's (1998) DPT suggests that, when antecedents are salient, anaphors that are immediately recognized as an anaphor (e.g., English overt pronouns) are read faster than anaphors that initially receive non-anaphoric interpretation. This is because the former type of anaphors leads readers to search for antecedents starting with salient ones and ending with non-salient ones while the latter type of anaphors leads readers to search for antecedents in an inverse situation, starting with non-salient ones and ending with salient ones. On the other hand, Almor's (1999) ILH explains that, when antecedents are salient, anaphors that carry little semantic information (e.g., English overt pronouns, Spanish null pronouns) are read faster than anaphors that carry rich semantic information. This is because the rich information is not necessary in order to identify a salient antecedent, which should be interpreted as an antecedent by default. On the other hand, when antecedents are not salient, since they are not the default antecedents, rich semantic information that an anaphor carries contributes to the identification of the non-salient antecedents. Thus, the advantage of anaphors with little information over anaphors with

rich information is decreased, nullified, (Gordon, Grosz & Gilliom, 1993) or even overturned (Gelormini-Lezama & Almor, 2011).³⁸

Returning back to the present study, Experiments 1 and 2 revealed that null pronouns and repeated names as well as null pronouns and overt pronouns exhibited the RNP and OPP, respectively, only when the overt anaphors were the non-topic anaphors marked by *ga*. The non-topic anaphors and null pronouns elicited the interactions with the antecedents' grammatical statuses, in a similar way to Gelormini-Lezama and Almor (2011) that found the RNP and OPP in Spanish. The RNP and OPP detected in Experiments 1 and 2 indicate that the ILH is applicable to Japanese, when overt anaphors do not carry the topic-hood. Because of the rich semantic information of the overt anaphors with *ga*, the processing of these anaphors was penalized for salient antecedents but justified to some extent for non-salient antecedents.

In contrast, with the topic-hood given by the topic-postposition *wa*, the overt anaphors and null pronouns were read similarly faster for the salient antecedents than for the non-salient antecedents, resulting in no RNP or OPP detected. It appears that the effect from topic-marking overrode the effect from semantic information that anaphors carry (i.e., the topic-marking to some extent neutralized the processing penalty from the rich semanticity that anaphors carry). Therefore, the DPT is applicable here instead of the

³⁸ Recall that, when antecedents were not salient, Gordon, Grosz and Gilliom (1993) found that sentences with English repeated names were read as fast as sentences with overt pronouns, and Gelormini-Lezama and Almor (2011) found that sentences with Spanish repeated names and overt pronouns were read faster than sentences with null pronouns.

ILH. When processing the topic anaphors with *wa*, readers immediately interpreted them as an anaphor and searched for the antecedent starting from salient antecedents to non-salient ones. Thus, the reading times of the topic anaphors with *wa* were faster when antecedents were salient than when they were not. This immediate anaphoric interpretation for the topic-*wa* anaphors contributed to the avoidance of the RNP and OPP.

Different from Experiments 1 and 2, Experiments 3 and 4 contrasted the topic-subject antecedents and non-topic-subject antecedents. Although the results of Experiments 3 and 4 were somewhat mixed, their main findings were that null pronouns and the repeated-name topic-*wa* anaphors elicited a marginal RNP, but repeated-name non-topic-*ga* did not. The results of Experiment 3 are therefore inconsistent with the results of Experiment 1, which showed that repeated-name topic-*wa* did not elicit an interaction while repeated-name non-topic *ga* did. In Experiment 3, an RNP with repeated-name topic-*wa* anaphors was elicited possibly because the difference of antecedent saliences (topic-subject vs. non-topic-subject) was subtle (compared to subject vs. object in Experiment 1). The relatively subtle difference of saliences was detected by null pronouns but not detected by repeated-name topic-*wa*, resulting in the marginal RNP, unlike the difference between subject antecedents and object antecedents in Experiment 1, which both null pronouns and repeated-name topic-*wa* detected (resulting in no RNP). In contrast, the repeated-name non-topic-*ga* did not show the RNP in Experiment 3 due to *ga*'s function to introduce new information as the subtopic. Since the non-topic-*ga* anaphors were interpreted as the subtopic which would be the most unlikely to be associated with the discourse topic, building a referential relationship between them

elicited long reading times. Since both the non-topic-*ga* anaphors and null pronouns dis-preferred the topic antecedents, they did not elicit the processing penalty (i.e., there was no interaction between these two types of anaphors and antecedents).

Experiments 5 and 6 utilized antecedents which differed in terms of empathy status. One finding in these experiments was the marginal OPP with overt pronouns regardless of topic-hood (when analyzing only the discourses that included *ageru/kureru*). This outcome is different from Experiments 1 and 2 in that the topic-hood of anaphors did not help readers avoid the processing penalty derived from the combination of salient antecedents (empathy locus) and semantically rich anaphors (overt pronouns), thus the DPT does not apply while the ILH is applicable. I attribute the disappearance of the *wa*'s facilitation to the anti-logophoricity that Japanese overt pronouns carry. This property may have lead readers to avoid the reference to empathy locus when encountering overt-pronoun anaphors. In Experiments 1 and 2, the empathy status of antecedents was not specified, so overt pronouns' anti-logophoricity was not active in readers' processing. Also, in Experiment 5, repeated-name anaphors do not carry anti-logophoricity, so these anaphors did not react to the empathy-status of antecedents.

The remaining question is why the analyses with the entire dataset (including the data from the discourses with *iku/kuru*-type of verbs) did not detect the OPP. The failure to elicit the OPP indicates that these verbs with *iku/kuru* did not provide the empathy-locus status to antecedents, which nullified the effect gained when the verbs were only *ageru/kureru*-type. One possibility is that, as briefly mentioned earlier, *iku* 'go' and *kuru* 'come' might lose the function to indicate empathy-loci when they are used as an auxiliary verb that is attached to another verb, like the items used in our experiments, i.e.,

motte-iku/motte-kuru ‘bring’ and *katte-iku/katte-kuru* ‘buy’. Another possibility is that *iku/kuru* are originally too weak in indicating empathy-loci, unlike *ageru/kureru* ‘give’-type of verbs. These possibilities should be tested in another experiment, and the experiment would need a large number of participants because, in the experiment, these verbs will have to be repeated several times in the exactly the same form, and thus there should not be many items in each condition.

The most noteworthy finding of this dissertation research was from Experiments 1 and 2. The overt topic anaphors marked by *wa*, both repeated name and overt pronoun, did not elicit the RNP or the OPP. In languages such as English or Spanish that do not explicitly mark the topic-hood, readers may initially interpret a subject anaphor merely as a grammatical subject, without interpreting it as the topic that should refer to already acknowledged information. In this situation, non-anaphoric interpretation (DPT) or redundant informativity (ILH) of the anaphor could penalize the processing of an anaphor. However, in a language such as Japanese that explicitly indicates the topic-hood of an anaphor, the effect from an explicit topic-marking may enhance the anaphoric interpretation and override the effect from the informativity of an anaphor. Gelormini-Lezama and Almor (2011) proposed that the RNP could be a universal phenomenon across languages, but the present study indicates that this argument should be confined to a condition in which the repeated names are not marked as the topic and thus not immediately interpreted as an anaphor. Also, Gelormini-Lezama and Almor suggested that the OPP should be observed in the null subject languages that utilize rich/informative verbal morphology, but the OPP can be extended to languages whose overt pronouns inherently connote rich semantic information such as those in Japanese.

Regarding the salience of antecedents, the present study, as well as earlier studies, showed robust evidences that the grammatical subjects are more salient than objects. Also, although I used a subset of the dataset (excluding *iku/kuru*), this study presented that empathy locus seems more salient than non-empathy locus. On the other hand, it may be unclear whether the topic or the non-topic is more salient than the other in a discourse. At the beginning of this dissertation, the term *salience* was defined as the high activation of an entity in the readers' working memory. However, this definition might be insufficient or inappropriate when examining the salience/accessibilities of the topic antecedents and the non-topic antecedents. In a Japanese discourse, a topic-*wa* entity remains activated as an overall topic during a reader's processing of the discourse, while a non-topic-*ga* entity is temporarily foregrounded (Maynard, 2004) and thus activated at the moment when readers process it and the surrounding sentences. In this situation, it is difficult to determine that one is more salient than the other in a definitive way because the relative saliences of entities fluctuates in the flow of a discourse, e.g., a non-topic could become temporarily more salient than the discourse topic. Accordingly, it is also difficult to argue that a short reading time of a sentence with an anaphor reflects the high degree of salience of its antecedent.

Another finding from Experiments 1-4 was that, in spite of their rich semantic information, sentences with Japanese overt pronouns tend to be read faster for salient antecedents than sentences with repeated names (unless their anti-logophoric property reacts to antecedents' empathy status as in Experiments 5 and 6). This is possibly because the overt pronouns are more likely to be interpreted as an anaphor than repeated names are, in accordance with the DPT. In Experiments 1 and 2, although both elicited the RNP

and OPP, overt-pronoun-*wa* were read significantly faster when antecedents were salient (subject) than when they were non-salient (object), while repeated-name-*wa* was not read differently between the antecedents at a significant level. In Experiments 3 and 4, the repeated-name-*ga* was read almost significantly faster when antecedents were the non-topic than when they were the topic, but the reading-time difference disappeared with the overt-pronoun-*ga*. These differences between repeated names and overt pronouns in terms of their sensitivities to antecedents could be because overt pronouns tended to receive anaphoric interpretation more than repeated names did, having led readers to search for the antecedents starting from salient ones.

Chapter 10. Conclusion

While Gordon, Grosz and Gilliom (1993) first reported the RNP, a following study by Gelormini-Lezama and Almor (2011) contributed to classifying between languages that obligatorily place an overt grammatical subject such as English and languages that utilize null subjects and rich verbal morphology like Spanish. A similar study on Italian and Brazilian Portuguese (Almor, de Carvalho Maia, Cunha Lima, Gelormini-Lezama, & Vernice, 2013) confirmed their analysis. The present study further develops the understanding of distinct effects between the above languages and a language that utilizes an explicit topic-marking morpheme, Japanese. This study newly classifies languages regarding the RNP and OPP where the RNP and the OPP are not elicited when a language marks anaphors as the topic, while the RNP and OPP are generated when a language does not mark anaphors as the topic. In order for this classification to be valid, further testing on other languages that utilize topic-marking morphemes is needed. Korean, which also allows distinct use of topic-subjects and non-topic-subjects, is a good candidate to be examined.

The second experiment of Gelormini-Lezama and Almor (2011) observed that the RNP and OPP disappeared when anaphors play the role of the discourse focus. On the other hand, the present study showed that, when an anaphor plays the role of the topic, it is not subject to the RNP or OPP. These results from the present study and Gelormini-Lezama and Almor imply that the use of repeated-name anaphors or overt-pronoun anaphors are not penalized when readers realize that these anaphors perform some kind

of discourse function, either as the topic or focus. This implication further predicts that, in a specific context-setting where Japanese subject anaphors with *ga* appear as the exhaustively-listed focus in critical sentences (i.e., the second sentences in the present study), we might not observe the RNP or the OPP, although Experiments 1 and 2 of the current study observed these effects on anaphors with *ga*, which were not what this study was focusing on.

This dissertation discussed that *wa* typically marks an NP that is previously acknowledged old information while *ga* marks a newly introduced NP (unless *ga* marks an exhaustively listed NP). However, according to Noda (1996), Kuno (1973a) and Makino (1996), an NP-*ga* can refer to an antecedent if sentences with the NP-*ga* describe unpredictable events in the discourse, e.g., ‘There was a friend-*ga* who was really good at math. ... One day, however, the friend-*ga* got zero on a math test (Makino, 1996, pp. 94-95)’. If this is the case, the *ga*-marked anaphor could be smoothly read when the sentence with the anaphor conveys unpredictable information, and the RNP and OPP detected with anaphors with *ga* in Experiments 1 and 2 might disappear. This would be an interesting topic for a future study.

To conclude, the present study revealed that the processing of referential expressions are greatly affected not only by the form of anaphors themselves, but also by the choice of postpositions, *wa* and *ga*, that mark the anaphors. There has been no study that has tested the RNP and OPP including the aspect of postpositions that mark anaphors. Therefore, this research fills a gap in cross-language understandings of processing referential expressions.

References

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psycholinguistic Review*, 106, 748-765.
- Almor, A., de Carvalho Maia, J., Cunha Lima, M. L., Gelormini-Lezama, C., & Vernice, M. (2013). *Universal mechanisms vs. language-specific patterns of use in processing null pronouns, overt pronouns and repeated names in Italian and Brazilian Portuguese*. Unpublished manuscript.
- Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clinton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. *Rivista di Linguistica*, 14, 1-19.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. New York: Routledge.
- Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperood & W. Spooren (Eds.), *Text representation* (pp. 29-88). Amsterdam: John Benjamins.
- Asayama, T. (2004). *Gendai Nihongo ni okeru “-wa” to “-ga” no imi to kinou* [Meanings and functions of wa and ga in contemporary Japanese]. Tokyo, Japan: Daiichi Shobo.
- Baayen, R. H. (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press. doi: 10.1017/CBO9780511801686
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278. doi: 10.1016/j.jml.2012.11.001

- Belletii, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language and Linguistic Theory*, 25, 657-689.
- Brown, S. R. (1970). *Colloquial Japanese: or, conversational sentences and dialogues in English and Japanese, together with an English-Japanese index to serve as a vocabulary, and an introduction on the grammatical structure of the language*. Tokyo, Japan: Hokushikan. (Original work published 1863, Shanghai: Presbyterian Mission Press)
- Chafe, W. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding* (pp. 21-52) Amsterdam: John Benjamins.
- Chamberlain, B. H. (1907). *A handbook of colloquial Japanese*. London: Trubner.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland: Foris Publications.
- Clancy, P. (1980). Referential choice in English and Japanese narrative discourse. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 127-202). Norwood, MA: Alber Publishing.
- Clancy, P., & Downing, P. (1987). The use of *wa* as a cohesion marker in Japanese oral narratives. In J. Hinds, S. K. Maynard & S. Iwasaki (Eds.), *Perspectives on topicalization: The case of Japanese wa* (pp. 3-56). Amsterdam: John Benjamins.
- Cloitre, M., & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes*, 3, 293-322.

- Colonna, S., Schimke, S., & Hemforth, B. (2012). Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics*, 50, 991–1013.
- Crawley, R. (1986). Some factors influencing the comprehension of pronouns. In C. Clifton (Ed.). *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Darnell, K., Boland, J., & Nakayama, M. (1994). The influence of orthography and sentence contrast on the processing of nouns in Japanese. *OSU Working Papers in Linguistics*, 44, 92-104.
- Deguchi, M. (2009). The contrastive reading of Japanese –wa, and the role of information structure. *Proceedings of the Arizona Linguistics Circle 3*, Tucson, AZ: University of Arizona.
- Dubinsky, S., & Hamilton, R. (1998). Epithets as antilogophoric pronouns. *Linguistic Inquiry*, 29, 685-693.
- Firbas, J. (1971). On the concept of communicative dynamism in the theory of functional sentence perspective. *Sbornik Praci Filosoficke Fakulty Brnenske University*, A19, 135-144.
- Firbas, J. (1979) A functional view of ‘ordo naturalis’. *Bruno Studies in English*, 13, 39-59.
- Fujii, N. (2005). Nihongo gakushusha no disukoosu ni okeru mondaiten [Problems of Japanese learners’ discourses]. In M. Minami (Ed.), *Gengogaku to nihongo kyoiku IV* [Linguistics and Japanese Education IV] (pp. 37-47). Tokyo, Japan: Kuroshio Publishers.

- Fukuda, Y. (2012). *Gengo shinrigaku nyumon* [Introduction to linguistic psychology]. Tokyo, Japan: Baifukan.
- Fry, J., & Kaufmann, S. (1998). Information packaging in Japanese. *Proceedings of the Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorical Grammar*. Saarbrücken, Germany: Universität des Saarlandes.
- Garrof, S., & Sanford, A. (1982). The mental representation of discourse in a focused memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics*, 35, 21-41.
- Garrof, S., & Sanford, A. (1988). Thematic subjecthood and cognitive constraints on discourse structure. *Journal of Pragmatics*, 12, 519-534.
- Gelormini-Lezama, C. (2010). *The overt pronoun penalty: A processing delay in Spanish anaphora comprehension* (Unpublished master's thesis). University of South Carolina, Columbia, SC.
- Gelormini-Lezama, C., & Amor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language Cognitive Processes*, 26, 437-454.
- Gelormini-Lezama, C. (2012). A review of the repeated name penalty: Implications for null subject languages. *Revista Linguistica*, 8, 22-34.
- Gelormini-Lezama, C., & Almor, A. (2014). Singular and plural pronominal reference in Spanish. *Journal of Psycholinguistic Research*, 43, 299-313.
- Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 5-41). Amsterdam: John Benjamins.

- Gordon, P. C., & Hendrick, R. (1997). Intuitive knowledge of linguistic coreference. *Cognition*, 62, 325-370.
- Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, 22, 389-424.
- Gordon, P. C., Hendrick, R., Ledoux, K., & Yang, C. L. (1999). Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes*, 14, 353-379.
- Gordon, P. C., Grosz, B. J., & Gillion, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311-347.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 41-58). New York: Academic Press.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21, 203-225.
- Guijarro, A. J. M. (2001). Topicality chains in two discourse genres. *Estudios Ingleses de la Universidad Complutense*, 9, 103-128.
- Gundel, J. K. (1988). *The role of topic and comment in linguistic theory*. New York: Garland Publishing.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English. *Journal of Linguistics*, 19, 377-417.
- Hagège, C. (1974). Les pronoms logophoriques. *Bulletin de la Société de Linguistique de Paris*, 69, 287-310.

- Hara, Y. (2006). Implicature unsuspendable: Japanese contrastive wa. *Proceedings of the 2004 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project.
- Hasegawa, N. (1995). Shoryakusareta daimeishi no kaishaku [The interpretation of omitted pronouns]. *Nihongogaku* [Study of Japanese], 14, 27-34.
- Hashimoto, S [Shinkichi]. (1948). *Kokugoho kenkyu* [Study of national language]. Tokyo, Japan. Iwanami Shoten.
- Hashimoto, S [Sachie]., Inui, K., Shirai, K., Tokubaga, K., & Tanaka, H. (2001). Nihongo-bun seisei ni okeru shouou hyougen no sentaku. [Choices of anaphoric expressions in composing Japanese sentences]. *Shizen Gengo Shori* [Natural Language Processing], 143, 33-40.
- Hasumi, S. (1977). *Han Nihongo ron* [Anti-Japanese language]. Tokyo, Japan: Chikuma Bunko.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In P. Porter & B. H. Partee (Eds.), *Meaning, use and the interpretation of language* (pp. 164-190). Berlin: Walter de Gruyter.
- Heycock, C. (2008). Japanese -wa, -ga, and information structure. In S. Miyagawa & S. Saito (Eds.), *The Oxford handbook of Japanese linguistics* (pp. 54-83). Oxford University Press.

- Hinds, J. (1975). Third person pronouns in Japanese. In F. C. C. Peng (Ed.), *Language in Japanese society* (pp. 129-157). Tokyo, Japan: University of Tokyo Press.
- Hinds, J. (1984). Topic maintenance in Japanese narratives and Japanese conversational interaction. *Discourse Processes*, 7, 465-482.
- Hinds, J. (1987). *Anaphora in discourse*. Edmonton, Canada: Linguistic Research.
- Hirotsu, M., & Schumacher, B. P. (2011). Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics*, 24, 274-292.
- Hoffmann, J. J. (2008). *A Japanese grammar*. Charleston, SC: Forgotten Books.
(Original work published 1868)
- Hirose, Y. (2000). Public and private self as two aspects of the speaker: A contrastive study of Japanese and English, *Journal of Pragmatics*, 32, 1623–1656.
- Hoji, H. (1991). Kare. In C. Georgopoulos & R. Ishihara (Eds.), *Interdisciplinary approaches to language: Essays in honor of Prof. S.-Y. Kuroda* (pp. 287-304). Reidel: Dordrecht.
- Hoji, H, Kinsui, S., Takubo, Y., & Ueyama, A. (2000). Demonstratives, bound variables, and reconstruction effects. *Generative linguistics in the old world (GLOW) in ASIA II*, 141–158. Nagoya, Japan: Nansan University.
- Iori, I. (1997). “Wa” to “ga” no sentaku ni kakawaru ichi youin: teijouhou meishiku no maaka no sentaku youin tono soukan karano kousatsu. [A factor for the choice wa and ga: Study of the relation with the choice of markers of definite noun phrases]. *Kokugogaku* [Study of national language], 188, 1-11.

- Iori, I. (2003). *“Zou wa hanaga nagai” nyumon: Nihongogaku no chichi, Mikami, Akira*
[Introduction to ‘zou wa hana ga nagai’: Akira Mikami, father of Japanese
linguistics]. Tokyo, Japan: Kuroshio Publishers.
- Ishiguro, K. (2013). *Nihongo wa “kuuki” ga kimeru* [Japanese language use, the
atmosphere decides it]. Tokyo, Japan: Kobunsha.
- Iwasaki, S. (1987). Identifiability, scope-setting, and the particle wa: A study of Japanese
spoken expository discourse. In J. Hinds, S. K. Maynard & S. Iwasaki (Eds.),
Perspectives on topicalization: The case of Japanese wa (pp. 83-106).
Amsterdam: John Benjamins.
- Iwasaki, S. (1993). *Subjectivity in grammar and discourse: Theoretical considerations
and a case study of Japanese spoken discourse*. Philadelphia: John Bnejamins.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge: MIT
Press.
- Jaeger, T. F., & Tily, H. (2010). On language ‘utility’: processing complexity and
communicative efficiency. *Cognitive Science*, 2, 323-335.
- Kameyama, M. (1985). *Zero anaphora: The case of Japanese*. (Doctoral dissertation),
Stanford University, CA.
- Kameyama, M. (1986). A property-sharing constraint in centering. *Proceedings of the
Twenty-fourth Annual Meeting on Association for Computational Linguistics
(ACL)*. Stroudsburg, PA: Association for Computational Linguistics.
- Kanaya, T. (2002). *Nihongo ni shugo wa iranai* [Subject is unnecessary in Japanese].
Tokyo, Japan: Kodansha.

- Kanaya, T. (2004). *Eigo ni mo shugo wa nakatta* [English also did not have subjects]. Tokyo, Japan: Kodansha.
- Kanno, K. (1997). The acquisition of null and overt pronouns in Japanese by English speakers. *Second Language Acquisition*, 5, 317-332.
- Kashiwadani, Y. (1984). Meishi, daimeishi no sho-mondai [Problems of nouns and pronouns]. In K. Suzuki & O. Hayashi (Eds.), *Kenkyu shiryō nihon bunpo I, hinshi: meishi, daimeishi* [Research resources on Japanese grammar I, parts of speech: nouns and pronouns], (pp. 188-218). Tokyo, Japan: Meiji Shoin.
- Kashiwagi, A., & Nakayama, M. (2008). NPs in L2 Japanese sentence processing. In T. Ogura, M. Hirakawa, H. Kobayashi, & H. Shirai (Eds.), *Studies in language sciences* 7 (pp. 169-185). Tokyo, Japan: Kuroshio Publishers.
- Kasuga, M. (1918). *Jinjou shogaku kokugo dokuhon no gohou kenkyu* [Linguistic study on elementary school textbook]. Tokyo, Japan. Shubunkan.
- Kawakami, S. (1996). *Ninchi gengogaku no kiso* [An introduction to cognitive linguistics]. Tokyo, Japan: Kenkyusha.
- Kaiser, E., Li, D. C.-H., & Holsinger, E. (2011). Exploring the lexical and acoustic consequences of referential predictability. In I. Hendrickx, A. Branco, S. Lalitha Devi & R. Mitkov. (Eds.), *Anaphora processing and applications*, (pp. 171-183) Heidelberg: Springer.
- Kinsui, S. (1989). Daimeishi to ninsho [pronouns and person-addressing terms]. In *Kooza nihongo to nihongo kyoiku IV, Nihongo no bunpo, buntai I* [Lectures on Japanese and teaching Japanese IV, Japanese grammar and styles I], (pp. 98-116) Tokyo, Japan: Meiji Shoin.

- Koga, Y. (2014). *Gendai Nihongo no “shiten” no taikai ni kansuru kenkyu* [A study on the system of “perspective” in contemporary Japanese]. (Unpublished doctoral dissertation). Kobe City University of Foreign Studies, Kobe, Japan.
- Kuno, S. (1972). Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry*, 3, 269-320.
- Kuno, S. (1973a). *Nihon bunpou kenkyuu* [Study of Japanese grammar]. Tokyo, Japan: Taishukan Shoten.
- Kuno, S. (1973b). *The structure of Japanese language*. Cambridge, MA: MIT Press.
- Kuno, S. (1976). Subject, theme and speaker’s empathy: A reexamination of relativization phenomena. In C. Li (Ed.), *Subject and topic* (pp. 417-444). New York: Academic Press.
- Kuno, S. (1977). Nihongo no shugo no tokushusei [Particularity of Japanese subject]. *Gengo* [Languages], 6, 11-18.
- Kuno, S. (1978). *Danwa no bunpou* [Grammar of discourse]. Tokyo, Japan: Taishukan.
- Kuno, S. (1986). Anaphora in Japanese. In S-Y. Kuroda, (Ed.), *Working papers from the first SDF workshop in Japanese syntax* (pp. 11-70). San Diego: UCSD Department of Linguistics.
- Kuno, S. (1987). *Functional syntax: Anaphora, discourse, and empathy*. Chicago: The University of Chicago Press.
- Kuno, S., & Kaburaki, E. (1977). Empathy and syntax. *Linguistic Inquiry*, 8, 627-672.
- Kuroda, S-Y. (1965/1979). *Generative grammatical studies in the Japanese language* (Doctoral dissertation). New York: Garland Publishing.

- Kuroda, S-Y. (1972). The categorical and the thetic judgment. *Foundations of Language*, 9, 153-185.
- Kuroda, S-Y. (1976). Nihongo no ronri, shikou [Logic and thinking in Japanese]. In S. Ono & T. Shibata (Eds.), *Nihongo I: Nihongogaku to kokugogaku* [Japanese I: Japanese and national language] (pp. 139-176). Tokyo, Japan: Iwanami Shoten.
- Kuroda, S-Y. (1992). *Japanese syntax and semantics*. Dordrecht, Holland: Kluwer Academic Publishers.
- Kuroda, S-Y. (2005). Focusing on the matter of topic: A study of wa and ga in Japanese. *Journal of East Asian Linguistics*, 14, 1-58.
- Kusano, K. (1978). Kokugo no tokuyuu seru gohou –soshu [Linguistic rule particular to Japanese –topic]. In S. Hattori, S. Ono, A. Sakakura & A. Matsumura (Eds.), *Nihon no gengogaku III: Bunpo I* [Japanese Linguistics III: Grammar I] (pp. 533-541). Tokyo, Japan: Taishukan Shoten. (Original work published 1899)
- Lambrechet, K. (1994). *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge University Press.
- Lane, D. M. (n. d.). *Log transformations*. Retrieved from:
<http://onlinestatbook.com/2/transformations/log.html>
- Larson, R., & Lujàn, M. (1989). *Empathic pronouns*. Retrieved from
<http://semlab5.sbs.sunysb.edu/~rlarson/emphpro.pdf>
- Li, C., & Thompson, S. (1976). Subject and topic: A new typology of language. In C. Li (Ed.), *Subject and topic* (pp. 457-489). New York: Academic Press.
- Makino, S. (1996). Uchi to soto no gengo-bunkagaku [Cultural linguistics regarding inside and outside]. Tokyo, Japan: Alc Press.

- Marslen-Wilson, W., Levy, W., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. Jarvella & W. Klein (Eds.), *Speech, place, and action* (pp.339–378). New York: Wiley.
- Martin, S. (1976). *A reference grammar of Japanese*. New Haven, CT: Yale University Press.
- Mathesius, V. (1981). *Obsahovy rozbor soucane anglictiny na zaklade obecne lingvisticke* [Functional analysis of English based on general linguistics]. (I. Iijima, Trans). Tokyo, Japan: Kiriara Shoten. (Original work published 1961)
- Matsumoto, K. (2006). *Sekaigengo he no shiza: Rekishi gongogaku to gengo ruikeiron* [A worldwide perspective on languages: Historical linguistics and linguistic typology]. Tokyo, Japan: Sanseido
- Matsushita, D. (1928). *Kaisen hyojun Nihongo bunpou* [Re-selected standard Japanese grammar]. Tokyo, Japan: Kigensha.
- Maynard, S. K. (1982). Analysis of cohesion: A study of the Japanese narrative. *Journal of Literary Semantics*, 11, 19-34.
- Maynard, S. K. (2004). *Danwa gongogaku: Nihongo no disukoosu o souzou suru kousei, retorikku sutorategii no kenkyu* [Discourse linguistics: A study of rhetoric strategy that build Japanese discourse]. Tokyo, Japan: Kuroshio Publishers.
- Mikami, A. (1960). *Zoo wa hana ga nagai* [Nose of an elephant is long]. Tokyo, Japan: Kuroshio Publishers.
- Mikami, A. (1963). *Nihongo no ronri: wa to ga* [Logic of Japanese: wa and ga]. Tokyo, Japan: Kuroshio Publishers.

- Mikami, S. (2012). *Two types of long scrambling in Japanese: A/A'-properties and the EPP on T in Focus-prominent languages*. Paper presented at Workshop on Altaic Formal Linguistics (WAFL), Stuttgart, Germany.
- Mio, I. (2003). *Mio Isago chosakushu* [Mio, Isago anthology]. Tokyo, Japan: Hitsuji Shobo. (Original work published 1948)
- Miyaji, A. (1964). Daimeisi [Pronouns]. In K. Morioka, M. Nagano, Y. Miyaji & T. Ichikawa (Eds.), *Kogo bunpo no mondaiten* [Problems of classic grammar] (pp. 200-209). Tokyo, Japan: Meiji Shoin.
- Murata, M. (1994). *Joji "wa" no subete* [All about the particle "wa"]. Tokyo, Japan: Shibundo.
- Nagano, M [Masaru]. (1972). *Bunshoron shosetu* [Comprehensive analysis of sentences]. Tokyo, Japan: Asakura Shoten.
- Nagano, M [Marisa]. (2012). *The use of overt pronouns by native and non-native speakers of Japanese*. Unpublished manuscript.
- Nagano, M [Marisa]. (2014). *Interpretation of overt pronouns in L1 & L2 Japanese: The role of context*. Poster session presented at the Second Language Research Forum (SLRF), Columbia, SC.
- Nair, V., & Almor, A. (2006). *On the measurement of the repeated name penalty with reading times and ERP*. Unpublished manuscript.
- Nakahama, Y. (2003). Development of referent management in L2 Japanese: A film retelling task. *Studies in Language and Culture*, 25, 127-146.
- Nakahama, Y. (2011). *Effects of task complexity, learners' L1 and proficiency level*. Tokyo, Japan: Hituzi Syobo.

- Nakayama, M. (1990a). *Priming and anaphoric expressions in Japanese*. Unpublished manuscript.
- Nakayama, M. (1990b). *Accessibility to the antecedents in Japanese sentence comprehension*. Unpublished manuscript.
- Nakanishi, K. (2001). Prosody and information structure in Japanese: A case study of topic marker *wa*. In N. Akatsuka & S. Strauss (Eds.), *Japanese/Korean Linguistics, 10*. (pp. 434-447). Stanford, CA: CSLI publications.
- Neeleman, A., & Reinhart, T. (1998). Scrambling and the PF interface. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 309–353). Stanford, California: CSLI Publications.
- Neeleman, A., Titov, E., van de Koot, H., & Vermeulen, R. (2009). A Syntactic Typology of Topic, Focus and Contrast. In J. van Craenenbroeck (Ed.), *Alternatives to cartography* (pp. 15-51). Berlin, Germany: Mouton de Gruyter.
- Newmeyer, F. J. (1999). Some remarks on the functionalist-formalist controversy in linguistics. In M. Darnell, E. A. Moravcsik, M. Noonan, F. J. Newmeyer & K. Wheatly (Eds.), *Functionalism and formalism in linguistics I: General papers* (pp. 469-486). Amsterdam: John Benjamins.
- Nishihara, S. (1990). Nichi-ei taisho shujiho. [Contrasting Japanese and English Rhetoric]. *Nihongo Kyoiku*, 72, 25-41.
- Nishimura, M. (1989). The topic-comment construction in Japanese-English code-switching. *World Englishes*, 8, 365–377.

- Niwa, T. (2004). Shugo to daimokugo. [Subject and Topic]. In K. Onoue (Ed.), *Asakura Nihongo kouza*. [Asakura Japanese lecture] (pp. 257-278). Tokyo, Japan: Asakura Shoten.
- Noda, H. (1996). *Wa to ga* [Wa and ga]. Tokyo, Japan: Kuroshio Publishers.
- Obana, Y. (2003a). Anaphoric choices in Japanese fictional novels: The discourse arrangement of noun phrases, zero and third person pronouns. *Text*, 23, 405-443.
- Obana, Y. (2003b). The use of kare/kanojo in Japanese society today. *New Zealand Journal of Asian Studies*, 5, 139-155.
- Ohkubo, Y. (1969). Saigo no hito-ha [The last leaf]. In *O. Henri Tanpenshu 3* [Anthology of O. Henry] (pp. 8-18). Tokyo: Shinchosha.
- Okamura, K. (1972). Daimeishi towa nani ka [What are pronouns?] In K. Suzuki & O. Hayashi (Eds.), *Hinshi-betsu nihon bunpo koza II: Meishi, daimeishi* [Parts of speech in Japanese grammar II: Nouns and pronouns] (pp. 81-121). Tokyo, Japan: Meiji Shoin.
- Okuma, T. (2011). Acquisition of topic shift by L1 English speakers of L2 Japanese. In J. Herschensohn & D. Tanner (Eds.), *Proceedings of the Eleventh Generative Approaches to Second Language Acquisition conference*. Somerville, MA: Cascadilla Proceedings Project.
- Okuma, T. (2014). *The OPC and task effects in interpreting Japanese pronouns*. Paper presented at the Second Language Research Forum (SLRF), Columbia, SC.
- Oshima, D. (2007a). On empathic and logophoric binding. *Research on Language and Computation*, 5, 19-35.

- Oshima, D. (2007b). Syntactic direction and obviation as empathy-based phenomena: A typological approach. *Linguistics: An Interdisciplinary Journal of the Language Sciences*. Retrieved from <http://www.thefreelibrary.com/Syntactic+direction+and+obviation+as+empathy-based+phenomena%3A+a...-a0168629406>
- Ono, S. (1976). Nihongo kenkyuu no rekishi II: Meiji ikou [History of the study of Japanese II: After Meiji period]. In S. Ono & T. Shibata (Ed.), *Nihongo I: Nihongo to kokugogaku* [Japanese I: Japanese and national language] (pp. 231-274). Tokyo, Japan: Iwanami Shoten,
- Ono, T., & Thompson, S. (2003). Japanese (w)atash / ore / boku 'I': they are not just pronouns. *Cognitive Linguistics*, 14, 321-347.
- Ono, T., Thompson, S., & Suzuki, R. (2000). The pragmatic nature of the so-called subject marker *ga* in Japanese: evidence from conversation. *Discourse Studies*, 2, 55-84.
- Onoe, K. (1977). Teidairon no isan [Legacy of discussion on topic]. *Gengo* [Languages], 6, 20-31.
- Palmroos, T. (2010). An 'I' for an 'I'... (Unpublished bachelor's thesis). Lund University, Lund, Sweden.
- Pimentel, C. L. (2014). *Pronominal interpretations in L2 Japanese*. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.
- Porter, S. (1948). The last leaf. In *The pocket book of O. Henry stories* (pp. 198-204). New York: Washington Square Press.

- Prince, E. F. (1978). On the function of existential presupposition in discourse. *Papers from the Fourteenth Regional Meeting*, Chicago, IL: Chicago Linguistics Society.
- Rohde, H., Kehler, A., & Elman, J. L. (2006). Event Structure and Discourse Coherence Biases in Pronoun Interpretation. *The Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society*, Vancouver, Canada: Cognitive Science Society.
- Saito, M., & Hoji, H. (1983). Weak crossover and move α in Japanese. *Natural Language and Linguistic Theory*, 1, 245–259.
- Sakuma, K. (1940). *Gendai nihongohou no kenkyu* [Study of contemporary Japanese grammar]. Tokyo, Japan: Kuroshio Publishers.
- Sanford, A. J., Moar, K., & Garrod, S. C. (1988). Proper names as controllers of discourse focus. *Language and Speech*, 31, 43-56.
- Sells, P. (1987). Aspects of logophoricity. *Linguistic Inquiry*, 18, 445-479.
- Shibatani, M. (1990). *The Languages of Japan*. Cambridge University Press.
- Shibatani, M. (1991). Grammaticization of topic into subject. In B. Heine & E. Traugotto (Eds.), *Grammaticalization II* (pp. 93-133). Amsterdam: John Benjamins.
- Shimojo, M. (2015). Discourse representation and argument linking: An analysis of zero anaphora and topicalization in Japanese narratives. In O. Hieda (Ed.), *Information structure and Nilotic languages: Studies in Nilotic linguistics, volume 10* (pp. 1-17). Tokyo, Japan: Tokyo University of Foreign Studies.
- Shoji, S., Dubinsky, S. & Almor, A. (2015). *Topics, subjects and pronouns*. Poster session presented at Neural Plasticity and Learning Conference. Columbia, SC.

- Shoji, S., Dubinsky, S. & Almor, A. (2016). *Topic-hood differently affects processing Japanese repeated names and pronouns*. Poster session presented at Annual CUNY Conference on Human Sentence Processing. Gainesvill, FL.
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22, 339-368.
- Suzuki, R., & Ono, T. (1991). Japanese *ga*, spotlighting, and intransitives in spoken narratives. *Santa Barbara Papers in Linguistics*, 3, 144-152.
- Suzuki, T. (1978). *Japanese and the Japanese: Words in culture*. Tokyo, Japan: Kodansha. (Original work published 1973 as *Kotoba to bunka* [Language and culture], Tokyo: Iwanami Shoten]
- Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience*, 16, 715 -726.
- Tanaka, K. (1989). *Kokkago o koete: Kokusaika no naka no nihongo* [Beyond national language: Japanese in globalism]. Tokyo, Japan: Chikuma Gakugei Bunko.
- Takubo, Y., & Kimura, H. (1992). Chugokugo, nihongo, eigo niokeru san-ninsho daimeishi no taisho kenkyu [Comparative studies of third person pronouns between Chinese, Japanese and English]. In Y. Okochi (Ed.), *Nihongo to chugokugo no taisho kenkyu ronbun-shu* [Essays on comparative studies between Japanese and Chinese] (pp. 137-152). Tokyo, Japan: Kurosio Shuppan.
- Tsuchiya, S., Yoshimura, N., & Nakayama, M. (2014). *Referentially dependent subject nouns in L2 Japanese narratives: a preliminary study*. Paper presented at the 2014 Annual Research Forum of the Linguistic Society of Hong Kong, Hong

- Kong. Abstract retrieved from
<http://www.lshk.org/sites/default/files/ARF%202014%20Program%20and%20Abstracts.pdf>
- Uehara, S. (1996). Are pronouns always zero in zero pronominal languages? The case in Japanese. In B-S. Park & J-B, Kim (Eds.), *Selected Papers from the Eleventh Pacific Asia Conference on Language, Information and Computation* (pp. 209-218). Seoul, Korea: Language Education and Research Institute, Kyung Hee University.
- Ueno, M., & Kehler, A. (2010). The interpretation of null and overt pronouns in Japanese: Grammatical and pragmatic factors. In S. Ohisson & R. Catrambone (Eds.), *Proceedings of the Thirty-second Annual Meeting of the Cognitive Science Society* (pp. 2057-2062). Austin, TX: Cognitive Science Society.
- Vermeulen, R. (2008). Topics, contrast and contrastive topics in Japanese. In R. Vermeulen & R. Shibagaki (Eds.), *MIT Working Papers in Linguistics*, 58 (pp. 361-372). Cambridge: MIT Press.
- Walker, M. A., Iida, M., & Cote S. (1994). Japanese discourse and the process of centering. *Computational Linguistics*, 20, 193-232.
- Walker, M. A. (1998). Centering, anaphora resolution, and discourse structure. In M. A. Walker, A. K. Joshi & E. Prince (Eds.). *Centering theory in discourse* (pp. 401-436). Oxford: Calarendon Press.
- Wang, L. & Schumacher, P. B. (2013). New is not always costly: Evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology*, 4, 1-20.

- Warnick, P. (1991). The use of personal pronouns in the language of learners of Japanese as a second language. *Deseret Language and Linguistic Society Symposium*, 17, 109-121.
- Weil, H. (1978). *The order of words in the ancient languages compared with that of the modern languages*. Amsterdam: John Benjamins. (Original work published 1869).
- Wetzel, P. J. (1994). A motivative self: the linguistic indexing of *uchi* and *soto*. In J. M. Bachnik & C. J. Quinn, Jr. (Eds.), *Situated meaning: Inside and outside in Japanese self, society, and language* (pp. 73-87). Princeton University Press.
- Wlodarczyk, A. (2005). From Japanese to General Linguistics: starting with the *wa* and *ga* particles. In A. Wlodarczyk, (Ed.), *Paris lectures on Japanese linguistics* (pp. 27-61). Tokyo, Japan: Kurosio Shuppan,
- Yamada, T. (1936). *Nihongo bunpogaku gairon* [Lectures on Japanese linguistics]. Tokyo, Japan: Saibunkan.
- Yanagimachi, T. (2000). JFL learners' referential-form choice in first-through third-person narratives. *Japanese Language Education around the Globe*, 10, 109-128.
- Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, 14, 715-743.
- Yashima, J. (2014). On the apparent unbindability of overt third-person pronouns in Japanese. *Natural Language & Linguist Theory*, 33, 1421-1438.
- Yoshida, E. (2005a). Discourse segment and Japanese referring expressions: Are these bare NPs or proper names? *Proceedings of the Nineteenth Pacific Asia Conference on Language, Information and Computation*. Taipei, Taiwan: PACLIC.

Yoshida, E. (2005b). Danwa rikai moderu kara mita nihongo meisi-ku no kaishaku ni tuite [On interpretation of Japanese noun phrases based on discourse comprehension model]. *Mie University Papers on Humanities*, 22, 129-140.

Yoshida, E. (2011). *Reference expressions in English and Japanese: Patterns of use in dialogue processing*. Amsterdam: John Benjamins.

Appendix A. Readings and Meanings of Words Excluded from Corpus Search of
kare (彼) ‘he’ and *kanojo* (彼女) ‘she’

Word	Reading	Literal meaning	Meaning of the word
何彼	nanika	‘what’ + ‘that’	various things
誰彼	darekare	‘who’ + ‘that’	various people
彼我	higa	‘he/she/they’ + ‘I’	he/she/them and I
彼岸	higan	‘that’ + ‘shore’	A Buddhist week during the Spring and Autumnal Equinox
彼比	karekore	‘that’ + ‘this’	various things
彼氏	karesi	‘he’ + ‘Mr.’	boyfriend
彼ら	karera	‘he’ + plural	they
彼等	karera	‘he’ + plural	they
彼女ら	kanojora	‘she’ + plural	them (specifically women)
彼女等	kanojora	‘she’ + plural	them (specifically women)
彼処	kasiko	‘that’ + ‘place’	over there
彼程	arehodo	‘that’ + ‘degree’	that much
彼奴	kayatu/kyatu	‘that’ + ‘person’	that person
彼方	kanata	‘that’ + ‘direction’	over there

彼某	karegasi/kagasi	‘that’ + ‘unknown’	that unknown person
元彼	motokare	‘former’ + ‘he’	ex-boyfriend
今彼	imakare	‘now’ + ‘he’	current boyfriend
前彼	maekare	‘previous’ + ‘he’	ex-boyfriend
彼所	asuko	‘that’ + ‘place’	over there
海彼	kaihi	‘sea’ + ‘that’	oversea
